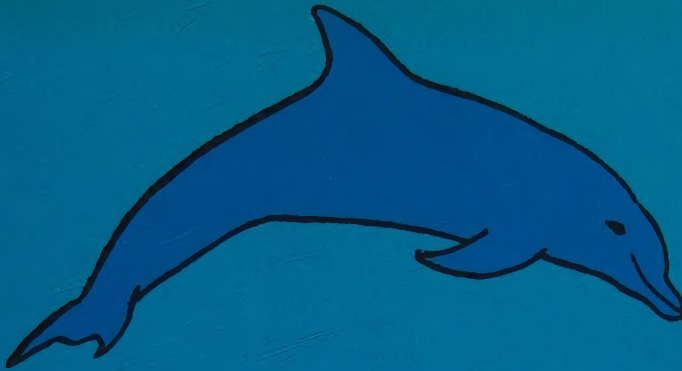


HIGHLY REPETITIVE DNA IN MARINE MAMMALS



BY

BENGT WIDEGREN

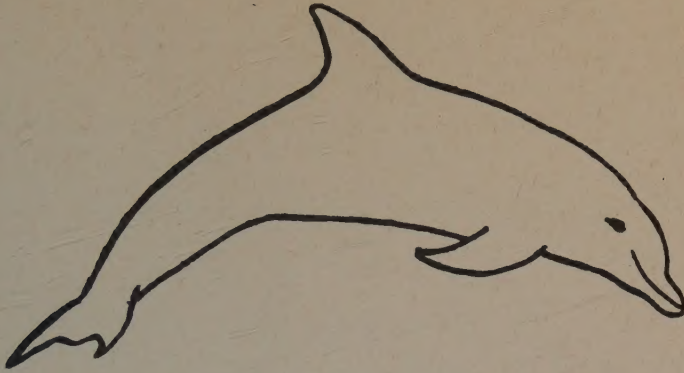


GENETISKA INSTITUTIONEN
Lunds universitet

DEPARTMENT OF GENETICS
University of Lund

1986

HIGHLY REPETITIVE DNA IN MARINE MAMMALS



BY

BENGT WIDEGREN



GENETISKA INSTITUTIONEN
Lunds universitet

DEPARTMENT OF GENETICS
University of Lund

*to my mother and
father*

List of papers

This thesis is based on the following papers, which are referred to by Roman numerals in the text.

- I. U. Arnason, M. Höglund and B. Widegren: Conservation of highly repetitive DNA in cetaceans. *Chromosoma (Berl)* (1984) 89:238-242.
- II U. Arnason and B. Widegren: Studies on ribosomal DNA in cetaceans (whales). *Hereditas* 101:149-154 (1984).
- III U. Arnason and B. Widegren: Different rates of divergence in highly repetitive DNA of cetaceans. *Hereditas* 101:171-177 (1984).
- IV B. Widegren, U. Arnason and G. Akusjärvi: Characteristics of a conserved 1579-bp highly repetitive component in the killer whale, *Orcinus orca*. *Mol. Biol. Evol.* 2(5): 411-419. 1985.
- V U. Arnason and B. Widegren: Pinniped phylogeny enlightened by molecular hybridizations using highly repetitive DNA. Submitted
- VI B. Widegren and U. Arnason: Analysis of the 420 bp highly repetitive component in balenopterid whales. Manuscript

INTRODUCTION

Highly repetitive DNA constitutes a substantial portion of the eucaryotic genomes. The amount may vary considerably and in some species highly repeated DNA may comprise as much as 80% of the total genome. In mammals the figure is in most cases 30 - 40% (Brutlag 1980, Singer 1982).

Repeated DNA occurs either interspersed or tandemly repeated and is commonly classified according to the number of repeats per genome. The classes are referred to as highly, moderately and low repetitive DNA. Until recently, repetitive DNA was considered to be limited to eucaryotes only, but recent findings show that repeated DNA, although at a low redundancy, is also present in procaryotes (e.g. Lin et al. 1984).

The possible functions of repetitive DNA have been discussed in a number of papers (e.g. John and Miklos 1979, Singer 1982, Jelenik and Schmid 1982, Jones and Singh 1982, Kao 1985) but the evidence for definite biological functions is still fragmentary. Interspersed repetitive DNAs such as Alu sequences currently evoke particular interest, whereas reports on possible functions of tandemly organized highly repetitive DNA are considerably fewer. Several authors (e.g. Orgel and Crick, 1980, Dolittle, 1982) have suggested that both interspersed and tandemly organized highly repetitive DNA is non-functional and have referred to it as "selfish DNA". They postulate that the amplification and conservation of the repetitive DNA is due to an ability (selfishness) of the sequences to be retained in the genome. However, these authors do not exclude the possibility of secondarily acquired biological functions.

Interspersed repetitive DNA

Eucaryotes in general are characterized by interspersed repetitive DNA, i.e. related DNA

sequences present in numerous sites in the genome, separated by unique DNA of variable length (Jelinek and Schmid 1982). The size of members of different classes of interspersed sequences varies but usually they fall into two groups, short interspersed repetitive DNA, 100-300 bp in length and long interspersed repetitive DNA with fragment lengths of 5000 - 7000 bp.

The long dispersed repeats are structurally similar to integrated retroviral DNAs and are flanked on each side by 300 - 600 bp repeats. In *Drosophila* and yeast it has been shown that the long dispersed repeats are mobile and frequently spread in the genome (Jelinek and Schmid 1982). The most plausible models for this process are based on reverse transcription and reciprocal recombination respectively (Rogers 1985, Temin 1985). In relation to the model based on reverse transcription, it is important to consider that long interspersed repetitive DNAs are transcribed by RNA polymerase II and that some are transcribed to such an extent that they can be detected as discrete bands on RNA gels. The transcribed RNAs are polyadenylated and contain 5' cap structure. The long dispersed repeats constitute a considerable portion of the eucaryote genomes, in *Drosophila* about 16 - 17% (Young 1979). The function of long dispersed repeats is unknown as is their origin and it is not clear if they have derived from retroviruses or if retroviruses have been derived from the long dispersed repeats. Transposons in bacteria show striking similarities to the long terminal repeats (Lewin 1985), for which reason the structure might be ancient.

The short interspersed repeats, described in several eucaryotes, are usually separated by unique sequences exceeding 1000 bp in length. The short interspersed repeats are transcribed into hnRNA but are more rarely seen among the polyA⁺ RNAs. Each individual repeat is flanked by small direct repeats at both ends. Homologies have been found between short interspersed repeats and tRNAs and U RNAs (Daniels and Deininger 1985). These findings may suggest that the mechanism for spreading this DNA involves reverse transcription of RNAs but the function, if any, of short interspersed repeats remains unknown.

The copy number of certain transcribed repeat families does not conform with the relative amount of repeats expressed as hnRNA, i.e. two families of interspersed repetitive DNA present in equal amounts in the genome can be transcribed to varying degrees. These findings have been used as an argument for a possible role in the regulation of transcription and processing of RNA, or in differentiation. The role in differentiation has been supported by findings that certain short period repeats show tissue specific transcription patterns. On the other hand it could be argued that tissue specific patterns are not the cause but the consequence of differential gene expression.

It is possible that Alu like sequences flanking a unique sequence may function as mobile elements for introducing DNA sequences at new chromosomal positions.

Tandemly repetitive DNA

Tandemly repeated DNAs occur in all eukaryotes and constitute a major portion of repetitive DNA. They are localized in heterochromatic regions of the chromosomes, in terms of transcription they are generally considered as being genetically inert. Usually tandemly highly repetitive DNAs are composed of short basic sequences, typically 2 - 12 bp in length. The basic sequence or sequences are then further repeated to give higher order units (Maio et al.1977). In restriction enzyme analysis, this organization is usually revealed by the presence of 2^{-n} meres in addition to multiples of the monomeric unit. Other tandemly organized repetitive DNAs may be of a considerable length without having internal repeats. In most cases, however, the length of the basic units is less than 1000 base pairs.

It is not clear how these tandem arrays evolved and by which mechanisms they spread in the genome. The unequal crossover model (Smith 1976) provides an explanation for the spreading and maintenance of the simpler repeats with relatively short basic sequences.

This model also satisfactorily explains the appearance of several related repetitive sequences in, for instance, *Drosophila*. A major weakness of this model is its inability to account for the existence of the same repeated sequences on nonhomologous chromosomes or at different positions on a single chromosome.

The tandemly highly repeated DNA is generally transcriptionally inert, but transcription has been detected in the lampbrush chromosome stage of oogenesis in, for example, the newt *Notophthalmus viridescens* (Stephenson et al. 1981). Transcription recorded in the lampbrush chromosomes is not necessarily functional and might only reflect a "run off" type of transcription at this stage of development.

Elimination of heterochromatic DNA takes place in polytene chromosomes of Diptera. Another interesting phenomenon is that described for *Ascaris* where one class of highly repetitive DNA is deleted in somatic cells but is retained in the germ line (Müller et al. 1982). These events have also been discussed in connection with the biological functions of highly repetitive DNA. These findings do not indicate, at least not for these classes of DNA, any biological function in somatic cells.

Functions of tandemly highly repetitive DNA are not known but some proposed effects of this type of DNA include position effects (variegation), involvement in chromosome pairing, effects on meiotic recombination, chromosome segregation, interaction of the chromosomes with the nuclear membrane and involvement in sex differentiation. In case any of these functions exist, they may have evolved secondarily after the proliferation of the component in the genome (Dolittle 1982). Genesis of a new class of highly repetitive DNA may affect chromosome pairing and meiotic recombination and consequently effect speciation (John and Miklos 1979).

It has been shown that DNA occurring as highly repetitive DNA sometimes amplifies into minute extra chromosomes and homogeneously staining regions (Simmons et al. 1984). In transformed cell lines, amplification of DNA may give rise to minute extra

chromosomes and to homogeneously staining regions. The amplification mechanism is not known in detail, but might be promoted by unequal crossover between sister chromatides (Maio et al. 1977, Stark and Wahl 1984).

Aims of the thesis

The aim of the present study was:

- I To obtain information of slowly evolving tandemly organized highly repetitive DNA in marine mammals.
- II To make use of highly repetitive DNAs for phylogenetic studies.

Cetaceans (whales) were chosen as the basic material for carrying out the study of highly repetitive DNA. This mammalian order is separated into two suborders, Odontoceti (toothed whales) and Mysticeti (whalebone whales), Fig. 1. Odontoceti comprise six extant families and Mysticeti three. The separation between the two suborders took place about 40 million years ago.

With the exception of the delphinid family, a class of tandemly organized highly repeated DNA, with a monomeric length of 1760 bp, was abundant in all extant cetacean families. In the delphinids, the 1760 bp component has been largely replaced by a 1580 bp unit, which differs from the 1760 bp component by a 180 bp deletion. This cetacean 1760 bp component has maintained its organization over a period of at least 40 million years as judged by Southern blot hybridizations, in situ hybridizations and sequencing studies (Arnason et al., 1982, paper I and IV).

Another cetacean highly repetitive component was also studied in some detail. Originally this component was referred to as a balenopterid heavy satellite (Arnason et al. 1978). Later it was cloned and sequenced as a 420 bp Xba I component (paper III and VI).

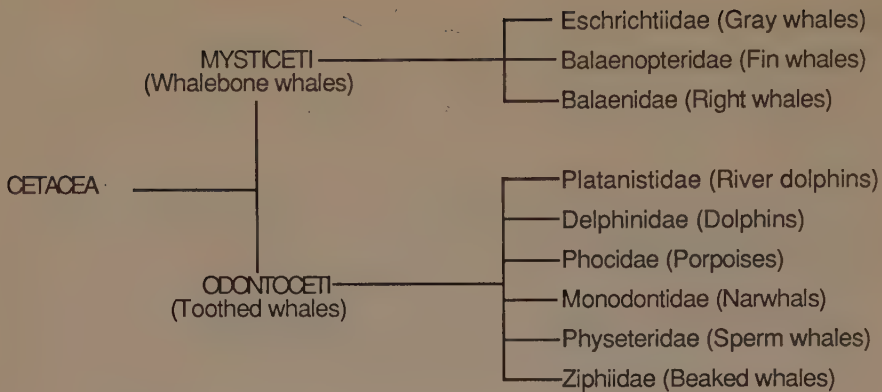


Figure 1. Schematic view of the order Cetacea with the two suborders Mysticeti and Odontoceti which separated approximately 40 million years ago.

A third cetacean highly repetitive component is present in centromeric localized heterochromatic regions of balenopterid chromosomes. This component does not exhibit the same rigid conservation of a basic unit as the two components discussed above. The evolution of this third component, which was originally isolated as a balenopterid light satellite (Arnason et al. 1978, paper III), appears to be much faster. The component did not hybridize to species outside the balenopterid family. It is conceivable that the age of this component is about 10 million years, the same as the evolutionary age of genus Balaenoptera.

In order to estimate the reliability of highly repetitive DNA as a tool for phylogenetic studies, the other large group of marine mammals, namely the pinnipeds, were studied. Different classes of highly repetitive DNA were used as probes to elucidate the pinniped phylogeny (paper V).

Materials and Methods

DNA and RNA isolation

Whenever possible, tissues, mostly liver and spleen, were collected from recently killed animals. The samples were subsequently frozen and stored at -20 or -70 °C. The DNA was prepared by the method developed by Marmur (1960) using RNase, pronase, α -amylase and phenol-chloroform treatments.

Plasmid or phage DNA were prepared using standard methods (Maniatis 1982).

RNA was isolated according to standard methods (Pritchard and Holland 1985) from freshly frozen material stored at -70 °C, or from fresh tissue cultures. The RNA was treated with DNase in the presence of RNasin before Northern blotting.

Sequencing

The Maxam-Gilbert sequencing was performed basically in accordance with the original protocol (Maxam-Gilbert 1980) with only minor modifications. Terminal transferase and α -³²P ddATP was used for 3' end labelling of protruding 3' ends or blunt ends. Klenow fragment together with some unique α -³²P dNTP was used for labelling recessed 3' ends. Calf intestine alkaline phosphatase, kinase and γ -³²P ATP were used for 5'-end labelling of blunt ends or protruding 5' ends. In some cases 3' end labelling was performed with the use of α -³²P ddNTP and Klenow fragment.

Dideoxy sequencing was performed according to Sanger (1981) after cloning in M13. Only shorter sequences of repetitive DNA were subjected to cloning in this vector. The sequencing was performed with α -³²P or α -³⁵S dNTP nucleotides.

The sequencing gels were standard, but wedge shaped, 0.1 - 0.4 mm in thickness. The gels were run at a constant temperature (55 - 60 °C) at a voltage of 1500 - 3000.

Southern and Northern blotting, hybridizations

Agarose gels (0.8 - 1.5%) were used for DNA and RNA. The filters used were nitrocellulose or nylon membranes. DNA, RNA and colony hybridizations were basically in accordance with Southern (1975) and Alwine (1980) using Denhardt's solution (Denhardt 1966) and denatured herring sperm or bacterial DNA to minimize background. The conditions used were mainly 3 - 6 x SSC, 64 °C during hybridization and 0.1 - 1 x SSC for washing. Autoradiography was performed using standard methods.

Cloning and transformation

Highly repetitive DNA's were isolated from preparative agarose gels (0.8-2%), electroeluted and purified on DEAE cellulose. The vectors used were pBR322, pAT153, pUC8/9 and M13 mp7/8/10/11. The bacterial strains used were HB101, MM294, GM169, JM83, JM101, JM103 or JM105. Transformation of the different bacterial strains was performed using standard protocols (Maniatis 1982). Positive clones were selected by means of colony hybridization.

Computer analysis

The analyses of the sequences were accomplished using three different computer program systems all run on a VAX 780 computer. The different program systems used were provided by Dr. Petter Gustafsson and Dr. Robert Harr of the Dept. of Microbiology, Umeå, Sweden, Dr. Philip Taylor, MRC Virology Unit, Glasgow, Scotland and the University of Wisconsin Genetics Computer Group (Deveraux et al. 1984) respectively. The data sequence library was provided by EMBL.

Tissue cultures, Chromosome preparations and in situ hybridization

Chromosome preparations were made from long term cell cultures established from

various tissues (lung, skin, cornea, retina, heart). C-banding was according to Sumner (1972). In the in situ hybridization experiments, the preparations were denatured in 0.1 x SSC just below boiling temperature (Singh et al. 1977). Hybridization was performed using nick translated cloned probes labelled with ^3H . Autoradiographic emulsion (Ilford) was used and the slides were exposed for 4 - 21 days.

RESULTS

The 1760 bp cetacean highly repetitive component

The 1760 bp cetacean highly repetitive component constitutes some 10-15% of the genome and is the most predominant of the highly repetitive DNA studied. The component is conspicuous after digestion with several restriction enzymes having a single recognition site within the repeat (see figures paper I).

Isolated or cloned fragments from different cetacean families were used as probes in several Southern blot hybridizations comprising representatives of all extant families, except the Eschrichtiidae, gray whales. These hybridizations revealed the presence of a monomeric unit, 1760 bp in length, in all families. Among the delphinids, however, the most prominent fragment length was 1580 bp and the 1760 bp unit was not so frequently represented. The relative distances between different restriction enzyme recognition sites was also shown to be conserved. The conserved localization of this component in the majority of C-bands in both baleen and toothed whales was verified by in situ hybridizations.

The base composition of the 1580 bp component from the killer whale is given in paper number IV and the composition of the 1760 bp component of the bowhead whale, aligned with the killer whale component, is given in Fig. 2. The 72 bp palindrome of the 1580 bp odontocete component (see paper IV) has been conserved in the 1760 bp mysticete component at the same position.

Transcription of the component was not detected in RNA from liver, spleen, pancreas, brain, testis, skin or cornea as judged from Northern blot hybridizations.

The component was compared with the EMBL sequence data library, version number 5, however, no noteworthy sequence homologies were found (paper IV). When the 1760 bp component was compared with version number 7 (released dec. 1985) some homologies were recorded. Two matches of possible interest were found with the human alphoid

satellite (Potter and Jones 1983) and with a mouse repetitive sequence (Martin et al. 1984), Fig. 3. In order to test the significance of these homologies statistically, the 1760 bp sequence was randomized, without changing the base frequencies, using the UWGCG program "shuffle". The constructed sequence was then compared to the sequence data library. Sequence homologies of almost the same magnitude were found.

Hybridizations to various mammals using the 1760 bp component as a labelled probe showed low degree hybridization to DNA of some deer species and dugong, but this has not been studied further.

The 420 bp balenopterid component

The 420 bp Xba I unit was also found to be tandemly repeated. The component is primarily located to distal heterochromatic regions of the mysticete karyotypes. The sequence is shown to be highly conserved among the balenopterids but also occurs in other whale families among both whalebone and toothed whales. However, in the toothed whales, the repeat length is different and the repetitivity lower. It is conceivable that this component evolved before the separation between toothed and whalebone whales took place. The 420 bp Xba I component also hybridized to a limited degree to other mammals and also to birds. Further results from the analyses of the 420 bp balenopterid component are presented in papers number III and VI.

The balenopterid light satellite (Eco RI component)

The balenopterids have a highly repetitive component, originally isolated as a light satellite by isopycnic ultracentrifugation. This component was later cloned and used for Southern blot and *in situ* hybridizations. The component is located in centromeric regions in a few chromosomes of the balenopterid karyotypes. The clustered organization of the component has not yet been resolved, but does not seem to be directly tandem repeated or to have definitive monomeric units in any of the species studied. Further results are reported in paper III.

	1		50
B.m.	ATCTTAGGAT	AAGTGTCTTT TCCTGTGTGA CTTATTTCAC TTAGAGTCAT	
O.o.	ATCTTATGAT	..GTTTCTTT TTTTGTGTGA CTTATTTCAT GTAGAATCAT	
	51		100
B.m.	.GTAGCTAAA	TCCACTCATC ATGCTGCTAA CGAGCCTTAT GACATTGATT	
O.o.	CGTACCTGAA	TCCACTCATT ATGCTGCT.A CGGGCCTGAT GACATAGATT	
	101		150
B.m.	TCATGGCTGA	GTGATATTCC ATTGTGCCTA AGTACCACAA CTTCTTGAT	
O.o.	TCATTGCTGA	GTGATATTGC ATTGTACGTA AGTACCACAA AGTCTT.TAT	
	151		200
B.m.	CCCTTTTTCT	CTTTCCTAGG ATATTGAAGG TGTACCGAGG TGGAGGTTCT	
O.o.	CCATTTTTCG	CTTT.CTGCG ATACTGAACT TGTACCATAA ACGAGGTTCT	
	201		250
B.m.	TGGAACACAGA	GCAGCCCTAA ACTTTGGGGT GCCTGTGTCT TGTGCGATTTT	
O.o.	TGTAACACAGA	GCCGTCACAA ACTTTGGGAT GGCTGTGTCT TTTTGATTTT	
	251		300
B.m.	TACTTTTCCC	AATTATACG CCCTAGAATG GAAGTGCCTT ATGCTCTCTA	
O.o.	AATTTCAACT	AAGCTATAGG ACCATAAGTG GAAGTGCCTT AGGCTCTGT.	
	301		350
B.m.	AGCTGTGTTT	ATTACATGTT TCGGGACACA CCGTACACTT CTCAGAGTG	
O.o.	TGCTTTGTG	TTTAGATGTT TCAGGAAACA CCATACACTT CTCAGAGTG	
	351		400
B.m.	GCTGTTGGCA	ATTTACATCC CGCCCATCAG CGTAGCAAGG CTCCCATGTC	
O.o.	GCTGTTGGCA	ATTTACATCC CGCCCATCAG CATAACAAGG CTCCCAATTC	
	401		450
B.m.	TCCATGGCCT	GTCCTGCITT TGCGGTTTTT ACACCTTTTTT CAGCATGGCC	
O.o.	TCCATGGCCT	GTCGCCGCTT TCTGGATTTT ACACCTTTTTT CAG.ATGGCC	
	451		500
B.m.	CTTTTGACCG	GTGGCAAGCG TGACTIONCTT TAGCGCTGAT TTGCCCTTCC	
O.o.	CTTTTGACCG	TGGGGCAGTG AGACTTCATG TAGTGCAGAT TTCCCTTTGCA	
	501	*****	550
B.m.	GGCTTGCTTG	GTTGGCCAAA AAGGGCGTTT GCGTTTTTTC CTGAATATAT	
O.o.	AGCTTGCTTG	GTTGGCCAAA AAGTTCGTAT GCGTTTCTTT CTGAATATAT	
	551	*****	600
B.m.	TCAGGAAAAA	ACGCAAAACGC CCTTTTGGC CAAGGGCATC ATTGTCAACG	
O.o.	TCAGGAAAAA	ACGCATAAGA CCTTTTGGC CAAGTGCATC ATTGTGGACG	
	601		650
B.m.	TTTGGCCGCT	TTTCATGTGC TTAAAGGCG AGTCGAATCT ACCTCCTGAA	
O.o.	TTCTGCCTGT	TTTCCTATGC TTACATGCA ATTCCAGTCT ACCTCCTGAA	
	651		700
B.m.	TTCTGTTTCC	TGTAATTCTG CCTTGCATTG AAGTCCTCTT CGTTGCGTTC	
O.o.	ATCGGTGTCT	GGCAATTTTG CCCGCTTTC AAGTCCTCTT GGCAGCCTTA	
	701		750
B.m.	CCTCAAGGTA	TTTGTGGACG GGAGTTATCA TTTATAACTC TGCAGGTTTG	
O.o.	CTTCAATATA	TTTTTGGACG ATAGCTGTCA TTTATAACTC TGCAGGTTTG	
	751		800
B.m.	TGAATTGCAG	TGCCCTGAG CTCCATTGT CAACTCGCTT TTTTGGGAGC	
O.o.	TGAATCACAG	TGCCCTGAG CTCCTTTCTT CAACTCGCTT TCTTGTGAGC	
	801		850
B.m.	TGGCCGCAAA	AGCGCAGGAC TGCTTCAGGC CCTATTCTGG TTCCGG..GG	
O.o.	TGGCTGCAAC	ACCGCAGGAT TGCTTGAGGC CCTCGTGTGG TTCCGGCAGG	
	851		900
B.m.	TCAGGCTGAG	CCTGTGGTTC ATTCTTCTTC CTGATGGGAA AATGGAGTGA	
O.o.	GCACGCTGAG	CCTTTTGTTA ATTCTTCTTC CTGGTGGGAA ATGAGGGTTA	
	901		950
B.m.	CCTGTGCTG	TCCCAACACC TAGGGCTAGT CTCCCATTGG TTCCCCCTCT	
O.o.	AATTTGCCCG	TCCAGACACC TCCAGCTAGT CTCTCATTGA TTCTCCCTAT	
	951		1000
B.m.	TCCCGTTCAT	CCTCAGGACA AATTGCCACC TGTACGAAAC AGGAGGTTGA	
O.o.	TCCTGTTCAT	TTTCCGCAGA AATTGCAAAC TGGGCCAAAC AGGAGGTTAA	

	1001					1050
B.m.	AGGTGC.GAT	TCTCCAAGTA	GGGAGATGGT	TAGTCAAGCG	GCTGGAATGG	
O.o.	AGGCACTGAC	TCTCCAAGTG	GGGAGAGTGT	TAGTAAAGCG	TCTGGAATGT	
	1051					1100
B.m.	CGAACCCGAG	CACAAGGAGA	GGAGAAATGA	GGTGCGTTTT	AGAAACCTTT	
O.o.	TGCACCCGAG	TACCAGGGGA	CGAAACTGA	GACACATTG	A.ACACGTTT	
	1101					1150
B.m.	CCCGATCACA	CGGTGTACCA	ACTTCTGGGT	TCCCCATGCC	TGCTTTCGAT	
O.o.	CCCGATCACA	CGGTGGATCA	TACTCTGGGT	TCCACATGCA	TGTTTTAGCT	
	1151					1200
B.m.	GTCCGAAGAT	GCCCTTGAAC	CTAGAGATT	GGGACCCATG	GAGTGGGTTC	
O.o.	GAAGGAAGAA	TCCCTTAAAC	CTGGAGAGTT	GAGACCCATG	GAATGGGTAC	
	1201					1250
B.m.	CCCGTGGTAT	TACTTTAAAG	GGTCTGCGTT	TGCT.ACCCA	ACCTCACCAC	
O.o.	CATGCAATAT	GACTTCAAAG	GGTCTGTATT	TGCTAACTGA	AACTCACTAA	
	1251					1300
B.m.	TCCTCTCAGC	CCCAAGAGTG	TCCAGCCTTA	TGTCTCATCC	AGCTGTGGGT	
O.o.	TCCTATCA..					
	1301					1350
B.m.	CTAGATATGG	TCAATCCAGG	TTGCATCAGA	GCCCTGAAC	GGATTTTGTA	
O.o.						
	1351					1400
B.m.	AGAATTTCTC	TCTCTTTCAC	TGTCCTTGTT	TGTGGGGTT	TTTTTTGAGT	
O.o.						
	1401					1450
B.m.	AATTTATGTT	TAGGATGAGG	CTTGGCTGCT	TTTCCCTGCT	GTGTTTATGC	
O.o.				CTGCT	GC GTTTATGC	
	1451					1500
B.m.	CGCTTTACCC	CCTCTTGATT	CACTTACAGA	GAGATATCAG	TACATAGGTT	
O.o.	CACTGTACAC	ACGCTTGATT	CCCTTTCGGA	GACATATAAG	TCCACAGGTT	
	1501					1550
B.m.	TTTCAAGTTCT	TACTAGTCAG	ATATACTCTT	GTGCGGTGAC	TAGGGGGTGT	
O.o.	TTAAGATTCT	TACTAGTCAG	GTATATTCTT	AGGAGTTTAA	TATGGAGTGT	
	1551					1600
B.m.	TGAGTCCAGT	TCACTGAGCA	GGGAGTAGGT	CTTGTCTATT	AGGTATTTGG	
O.o.	TGAGTCCA.C	TCGATCAGCA	AGGAGTAGCT	CTTGTCTATT	AAATATTTGG	
	1601					1650
B.m.	CTTATGGAAC	GGTCTCTGAG	CTAATTTCAC	ACCCTTGTTT	TAGGCATCAC	
O.o.	CTTATGGAAC	GGTATCTGTG	CTGATTTCAC	TCTCTGGTTT	TATGCAGCAC	
	1651					1700
B.m.	CCCACCTCAC	CTTTCCCGTT	AAGCAGCCAT	.CGTGTGTCT	TCTAAAGGTG	
O.o.	CTCAACTCAC	CTTTCCCGTT	AAGCAAGCAT	AAGTTGGTGT	TCTACATTTG	
	1701					1750
B.m.	TGACTCTGTT	CTATTTTGTA	ATTCAGTTCA	AGAGTAGCGA	TTTTGACATT	
O.o.	AGACCCTGTT	CTGTTTTGTA	ATTCAGTTCC	TGTGTAGCCA	AGTTTACATT	
	1750	1765				
B.m.	CCCTCTATAA	GAGAT				
O.o.	CCGTGTAGTA	GTGAT				

Figure 2. Alignment of the 1760 bp component of the right whale (B.m.) and the 1580 bp component of the killer whale (O.o.). The dots symbolize gaps introduced at the alignment. The large gap from position 1259 to 1436 represents the 177 bp deletion of the 1760 bp component producing the 1580 bp delphinid component. The 72 bp palindrome in positions 512 - 583 is marked by asterisks (*) .

gb2106.10 check: 3893 from: 518 to: 1760
mmllmd4.mbl check: 9759 from: 1 to: 984

Gaps: 14 Quality: 60.4 Ratio: 0.124 Words: 18 Width: 5 Limits:

+/-6

```

982 GGAAGGGCAAATCAGCGCTAAA.....AGAAGTCACGCTTGCCACCGGT 1025
    IIII IIIIIII III II I I I I IIII II
465 GGAAATGCAAATCA.....AAACAACCTGAGATTCCACCTCACACCAGT 509

1026 CAAAAGGGCCATGCTGAAAA...AAGTGTAACCAAGCAAAGCAGGAC 1071
    II II III I I I IIII I III II III I II II
510 CAGAATGGCTAAGATCAAAAATTCAGGTG.....ACAGCAGATGCTGGCG 554

1072 AGGCCATGGAGACATGGGAGCCTTGCTACGCTGATGGGCGGGATGTAAAT 1121
    III IIIII I III I I II I II I II IIIII I
555 AGGATGTGGAGAAGAGGAACACTCCTCCATTG.TTGGTGGGATTGCAGG 603

1122 TGCCACAGCCACTCTGGAGAAGTGTACGGTGTGTCCCGAAACATGTAAT 1171
    III IIIIIIIII I II II I III I I I I I
604 CTGTACAACCACCTCTGGAATCAGTCTGGCGGTTCTCTCAGAAA...ATT 650

1172 AAACACAG..CTTAGAGAGCATAGGGCACTTCCATTCTAGGGCGTATA.. 1217
    III II II III II III I II II IIII IIII
651 GGACATAGTACTACCGGAGGATCCAGCAATACCTCTCCTGGGCATATATC 700

1218 .....AATGGGAAAAGTAAAAATCGACAAGACACAGGCACCCCAAAGT 1261
    II II I II I I IIIII II II I I II
701 CAGAAGAAGCCCCAACTGGTAAGA.....AGGACACATGCTCCACTATGT 745

1262 TTAGGGCTGCTCTGTTTCCAAGAACCTCCACCTCGGTACACCTTCAATAT 1311
    I I II II I III II I II I II I I I I I I II
746 TCATAGCAGCCTTATTTATAATAGCCAGAACTGGAAAGAAC..CCAGAT 793

1312 CCTAGGAAAGAGAAAAAGGGATACAAGAAGTTGTGGTACTTAGGCACAAT 1361
    I II III II IIIII IIII IIIIIII I IIIII
794 GCCCCTCAACAGAGGAATGGATAC.AGAAAATGTGGTACATCTACACAAT 842

1362 GGAATATCACTCAGCCATGA.....AATCAATGTCATAAGGCTCGTTAGC 1406
    III II IIIIIII II I III III I II II
843 GGAGTACTACTCAGCTATTTAAAAAGAATGAATTTATGAA.....ATTCT 887

1407 AGCATGATGAGTGGATTAGTCTACGATGACTCTAAGTGAAATA....AGT 1452
    III III IIII I I I II I II IIIII II I I
888 AGCCAAATGGATGGACCTGGAGAGCATCATCCTGAGTGAGGTAACACAAT 937

1453 CACACAGGAAAAGACAC 1469
    IIII IIIII IIII
938 CACAAAGGAACCTCACAC 954
```



```

gb2106.10  check: 3893  from: 1      to: 1760  /reverse
           hssat03.mbl  check: 2525  from: 466   to: 1689
Gaps: 14   Quality: 58.2  Ratio: 0.109  Words: 12  Width: 7  Limits:
+/-8

1506 TTTTGACATTCCCTCTATAAGAGATATCTTAGGATAAGTGTCTTTTCCTG 1457
    | | | | | | | | | | | | | | | | | | | | | |
721  TGTTC AATTCCCACCTATGAGTGAGAACATGTGGTGTGTTGGTTTTTGTGTC 770

1456 TGTGACTTATTTTCACTTAGAGTCATCGTAGCTAAATCCACTCATCATGCT 1407
    | | | | | | | | | | | | | | | | | | | | | |
771  CTTGTGATAGTTTGCTGAGAATGATGGTTTCCAGCTTCATCCATGTCCT 820

1406 GCTAACGAGCCTTATGA.....CATTGATTTCATGGCTGAGTGATATTC 1363
    | | | | | | | | | | | | | | | | | | | | | |
821  GCAAAGGA....CATGA ACTCATCAT..TTTTATGGCTGCATAGTATTC 864

1362 CATTGTGCCTAAGTACCACAACCTCTTGATCCCTTTTTCTCTTTCTAG 1313
    | | | | | | | | | | | | | | | | | | | | | |
865  CATGGTGTATATGTGCCACATTTCT..TAATCCAGTCTATCATT.GTTG 911

1312 GATATTGAAGGTGTACCGAGGTGGAGGTTCTTGGAACAGAGCAGCCCTA 1263
    | | | | | | | | | | | | | | | | | | | | | |
912  GACATTTGGGTTGGTTCCAAGTCTTGCTTTTGTGAATAATGCCACAATA 961

1262 AACTTTGGGGTGCCTGTGTCTT.....GTCGATTTTACTTTTCCCA 1221
    | | | | | | | | | | | | | | | | | | | | | |
962  AACATACGTGTGCATGTGTCTTTATAGCAGTATGATTATAGTCCTTTGG 1011

1220 ATTTTATACGCCCTAGAATGGAAGTGCCCTATGCTCTCTAAGCTGTGTTTA 1171
    | | | | | | | | | | | | | | | | | | | | | |
1012 GTATATACTCAGT..AATGGGATGGCTGGGT.....CAAATGGTATTTTC 1053

1170 T....TACATGTTTCGGGACACACCGTACACTTCTCCAGAGTGGCTGTT 1126
    | | | | | | | | | | | | | | | | | | | | | |
1054 TAGTTCTAGATCCCTGAGGAATCGCCACACTGACTTCCACAATGGTTG.. 1101

1125 GGCAA....TTTACATCCCGCCCATCAGCGTAGCAAGGCTCCCATGTCTC 1080
    | | | | | | | | | | | | | | | | | | | | | |
1102 ...AACTAGTTTACAGTCCCACCAACAGTGTAAGTGTTCCTATTTCTC 1148

1079 CATGGCCTGTCTGCTTTTTCGGGTTTTTACACTTTTTTCAGCATGGCCCT 1030
    | | | | | | | | | | | | | | | | | | | | | |
1149 CACATCCTCTCCAGCACCTGTGTTTCCTGAC.TTTTAAATGATTGCCAT 1197

1029 TTTGACCGGTG.GCAAGCGTGACTTCTTTTAGCGCTGATTGCGCCTTCGG 981
    | | | | | | | | | | | | | | | | | | | | | |
1198 TCTAACTGGTGTGAGATGGTATCTCATGTGGTTTTGATTTC..... 1240

980  GCTTGCTTGGTTGGCCA 964
    | | | | | | | | | |
1241 .ATTTCCTGATGGCCA 1256

```

Figure 3. Two alignments between the 1760 bp component and an unidentified reading frame of the mouse L1 interspersed repeat (a) and between the human alphoid satellite(b). The gap weight in the comparisons is 5.00 and the length weight is 0.00.

Ribosomal DNA

Studies of polymorphism in rDNA were performed in three different species, namely the fin, sei and sperm whales. The observed polymorphism in these species appeared to be of the same magnitude as that among terrestrial mammals (paper number II).

Highly repetitive DNA and pinniped phylogeny

The occurrence of 4 unrelated highly repetitive components of true seals was studied. With respect to repeat lengths, all four randomly selected components were preserved within the pinnipeds, represented by true seals (Phocidae), sea lions (Otariidae) and the walrus (Odobenidae). Hybridization also occurred to terrestrial carnivores, notably the mustelids. The results from this phylogenetic study are reported in paper number V.

Discussion

The stability of the 1760 (1580) bp cetacean highly repetitive DNA and the 420 bp balenopterid Xba I component is striking. Four, randomly selected, pinniped highly repetitive DNA components also showed prominent conservation within the pinnipeds and also had the same repeat lengths in the mustelids (paper V).

Marine mammals in general are characterized by a conspicuous karyotype uniformity. Only two chromosome numbers, $2n=44$ and $2n=42$, occur among cetaceans. The factors considered to be primarily responsible for the slow karyotype evolution in marine mammals are (1) low productivity (late sexual maturity, not more than one offspring per year), (2) good vagility, and (3) an environment without distinct physical barriers (Arnason 1972).

The following questions can be raised with respect to the slow evolution of some of the components studied; (1) Is the conservation of highly repetitive DNA a particular characteristic of marine mammals and (2) have the marine mammals a slower mutational rate than terrestrial mammals or a more efficient repair system?

The first question can be answered negatively. Since the earliest studies of highly repeated DNA, based on components isolated by ultracentrifugation, it has been a common notion that this part of the genome undergoes rapid evolution and on these grounds, highly repetitive DNAs have been considered to be species specific (Brutlag 1980). Today the validity of this view is doubtful since many reports have demonstrated high conservatism among the highly repetitive DNA, of various organisms (e.g Lima de Faria et al. 1984, Jones and Singh 1982).

In addition to the evolutionarily conservative repetitive components mentioned above, an additional component was studied in the genus Balenoptera. This genus is about 7 - 10 million years old. This additional component, the so called light DNA satellite, had a species specific restriction pattern. This satellite is conceivably younger than the two other highly repetitive cetacean components studied which are found at higher taxonomic

levels. Nevertheless, the diversity of the light satellite was found to be greater than that of the two other components (paper III). The degree of polymorphism in cetacean rDNA does not seem to differ from that of other mammals (paper II). From these findings it can be concluded that the cetaceans do not have an overall molecular evolution that is slower than that of other mammals.

The cetacean 1760 (1580) bp component, which occurs in great numbers in all cetacean families, is interesting because of its striking conservation and because of the length of the basic repeat unit. It can only be conjectured how components such as the 1760 (1580) bp component can maintain fragment length and base composition so stringently over a period of at least 40 million years. From the different models of amplification of DNA mentioned in the introduction, the rolling circle model (Maio et al. 1977) appears to be the most plausible. The model of unequal crossing over (Smith 1976) would presumably generate too many alterations in a repeat of such basic length. Moreover, it can hardly explain the occurrence of this repeat in most of the C-bands present in the cetacean karyotypes. Proliferation and spreading via reverse transcription is an appealing model for interspersed repeats and may provide an explanation for the first events in forming an extra chromosomal circular molecule of the repeat. Another possibility in the formation of a circular molecule, is endogeneous restriction of the tandem repeat at the position of the large palindrome in the repeat (Fig. 2). Transcription of the component was not recorded in any of the tissues studied (immature testis, pancreas, brain, liver, spleen, skin and cornea) but the possibility of transcription in germ line and early embryonic development is not excluded. The component had several open reading frames but in the components analysed, the maximum length of these was about 330 bp. These findings do not exclude a possible presence of open reading frames covering the full length of the repeat among other members of the component. When the pros et cons of the various possibilities are considered, the model of the rolling circle is thought to be the most plausible for the spreading of this component since it, more easily than the other models, can account for the conservation of the sequence and its occurrence at many different positions in the karyotype.

The question how fragment length and base composition of the component has been maintained is also intriguing. It has now been well established that the fragment length of the component is different in the delphinid family in comparison with other cetacean families, although the 1760 bp repeat unit still exists in low numbers in the different genomes. This indicates that the entire length of the component is not of fundamental importance to the organism and/or for the repeat. This argumentation, however, might be flawed if the component has a biological function and this function can be accomplished by a few copies of the repeat. Biological functions determined by the base composition of the entire repeat are, however, not thought to be likely.

The striking preservation of the 1760 (1580) bp component indicates that its presence in the cetacean genome is strongly favoured. It is quite possible that sequences, generated in the genomes as rare events can adapt to act in a parasitic fashion and thus conserve their structure with respect to both length and base composition. This would fit the model of "selfish DNA" postulated by Orgel and Crick (1980). Provided the parasitic DNA secondarily acquired properties are advantageous for the organism, the possibilities of "selfish DNA" to be maintained in the genome would of course be increased. Such a situation might be considered as a symbiotic relationship. Advantages for the organism in such a symbiotic relationship might relate to the biological effects of the presence of heterochromatin, such as effects on crossover frequencies, and position effects. Other effects might involve chromosome pairing and segregation, and attachment of chromosomes to the nuclear membrane.

From the present knowledge of biological processes at the level of the cell and organism it is difficult to consider any biological function evolving from the base composition of tandemly repeated highly repetitive DNA, as it is unlikely that any organism would need such an amount of a specific sequence for the maintenance of a specific biological function. Therefore it is more appealing to look for possible biological effects, by the presence of many members of the same family at specific positions in the genome.

The conclusions reached here are that the amplification of the 1760 bp repeat is most easily explained by the model of the rolling circle and that the component does not have

any inherent biological function, but may act in a symbiotic way to give effects positive to the genome by an accumulation at specific positions.

The age of the 1760 (1580) bp component must exceed 40 million years, i.e. the time of the separation of odontocetes and mysticetes. Provided the homologies with the mouse repetitive DNA and the human alphoid satellite (Fig. 3), found by computer analysis are not due to coincidence, it is conceivable that the component was present in the early mammalian evolution. However, at the present stage the occurrence of these homologies, does not permit discussion based on possible functions.

The evolutionary stability of the two cetacean highly repeated DNA components and the four pinniped highly repeated components, clearly demonstrate that comparisons between highly repeated DNAs can be used to answer phylogenetic questions. These findings also support further research on the phylogenetic origin of highly repeated DNAs in order to explain their appearance and their possible biological functions in the genomes.

ACKNOWLEDGEMENTS

I wish to express my deepest gratitude to my supervisor and friend Dr. Ulfur Arnason, who introduced me to the field of molecular genetics, for his continuous support and encouragement, and many stimulating discussions during my doctoral work. I also want to express my sincere thanks to Mrs Anette Gullberg, Elinor Johansson and Margareta Sunner for their skilful help in the laboratory, to Dr. Ronald Pero for fruitful discussions on DNA chemistry and "Missouri squeeze" and to Professor Antonio Lima de Faria for encouragement at various stages of my work. The valuable collaboration with Dr. Göran Akusjärvi is also acknowledged as well as the numerous discussions with my friends Drs. Ulf Wiberg and Nils Mandahl. The support of all friends and colleagues at the Wallenberg laboratory and at the Department of Genetics is also acknowledged as are my thanks to Prof. Jan-Eric Edström for various suggestions at the preparation of the thesis and to Mrs Penny Kenton-Russ for her qualified work on the text. I also want to thank my family for continuous interest and encouragement.

The work has been generously supported by grants from the Erik Philip- Sörensen Foundation, the Nilsson-Ehle Foundation and the Helge Ax:son Johnsson Foundation.

Literature cited

Alwine, J.C., Kemp, D.J., Parker, B.A., Reiser, J., Renart, J., Stark, G.R. and Wahl, G.M., Detection of specific RNAs or specific fragments of DNA by fractionation in gels and transfer to diazobenzoyloxymethyl paper. *Meth. Enzym.*, 68, 220, 1980.

Arnason, U., The role of chromosomal rearrangement in mammalian speciation with special reference to Cetacea and Pinnipedia. *Hereditas*, 70, 113-118, 1972.

Arnason, U., Purdom, I.F. and Jones, K.W., Conservation and chromosomal localization of DNA satellites in balenopterid whales. *Chromosoma*, 66, 141-159, 1978.

Arnason, U., Purdom, I.F. and Jones K.W., Cetacean molecular hybridization using balenopterid satellite DNA cRNAs as probes. *Hereditas* 97, 33-36, 1982.

Blackburn, E.H. and Challoner, P.B., Identification of a telomeric DNA sequence in *Trypanosoma brucei*. *Cell*, 36, 447-457, 1984.

Brutlag, D.L., Molecular arrangement and evolution of heterochromatic DNA. *Ann. Rev. Genet.* 14, 121-144, 1980.

Daniels, G.R. and Deininger, P.L., Repeat sequence families derived from mammalian tRNA genes. *Nature*, 317, 819-822, 1985.

Denhardt, D.T., A membrane filter technique for the detection of complementary DNA. *Biophys. Res. Commun.*, 23, 641-646, 1966.

Deveraux, J., Haeberli, P. and Smithies, O., A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acid Res.* 12, 387-395, 1984.

Dolittle, W.F., Selfish DNA after fourteen months. *Genome evolution*, eds. Dover and Flavell, Academic Press (London) 1982.

Jelinek, W.R. and Schmid, C.W., Repetitive sequences in eukaryotic DNA and their expression. *Ann. Rev. Biochem.* 51, 813-844, 1982.

John, B. and Miklos, G.L.G., Functional aspects of satellite DNA and heterochromatin. *Int. Rev. of Cytol.* 58, 1-114, 1979.

Jones, K.W. and Singh, L., Conserved sex-associated repeated DNA sequences in vertebrates. *Genome evolution (monogr.)* ed. Dover and Flavell, 135-154, 1982.

Kao, F-T., Human genome structure. Int. Rev. Cytol., 96, 51-88, 1985.

Lewin, B., Genes II, John Wiley & Sons, 1985.

Lima de Faria, A., Arnason, U., Widegren, B., Essen-Möller, J., Isaksson, M., Olsson, E., and Jaworska, H., Conservation of repetitive DNA sequences in Deer species studied by Southern blot transfer. J. Mol. Evol. 20, 17-24, 1984.

Lin, R-J., Capage, M. and Hill, C.W., A repetitive DNA sequence, rhs, responsible for duplications within the *Escherichia coli* K-12 chromosome. J. Mol. Biol. 177, 1-18, 1984.

Maio, J.J., Brown, F.L. and Musich, P.R., Subunit structure of chromatin of eukaryotic highly repetitive DNA: Recurrent periodicities and models for the evolutionary origins of repetitive DNA. J. Mol. Biol. 117, 637-655, 1977.

Maniatis, T., Fritsch, E.F. and Sambrook, J., Molecular Cloning. Cold Spring Harbor Laboratory. 1982.

Marmur, J., A procedure for isolation of deoxyribonucleic acid from microorganisms. J. Mol. Biol., 3, 208-218, 1961.

Martin, S.L., Voliva, C.F., Buton, F.H., Edgell, M.H. and Hutchison, C.A., A large interspersed repeat found in mouse DNA containing a long open reading frame that evolves as if it encodes a protein. Proc. Natl. Acad. Sci. (USA) 81, 2308-2312, 1984.

Maxam, A.M. and Gilbert, W., Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol., 65, 499-560, 1980.

Müller, F., Walker, P., Aeby, P., Neuhaus, H., Back, E. and Tobler, H., Molecular cloning and sequence analysis of highly repetitive DNA sequences contained in the eliminated genome of *Ascaris lumbricoides*. Embryonic Development, part A, 127-138, 1982.

Orgel, L.E. and Crick, F.H.C., Selfish DNA: the ultimate parasite. Nature, 284, 604-607, 1980.

Potter, S.S. and Jones, R.S., Unusual domains of human alphoid satellite DNA with contiguous non-satellite sequences: sequence analysis of a junction region. Nucl. Acids. Res. 11, 3137-3153, 1983.

Pritchard, R.H. and Holland, I.B.(ed), Basic Cloning Techniques. (Monogr.) Blackwell Scientific Publications, 6-9, 1985.

Rigby, P.W., Dieckmann, M., Rhodes, C. and Berg, P., Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol., 113, 236-251, 1977.

Rogers, J.H., The origin and evolution of retroposons. *Int. Rev. of Cytol.* 93, 188-279, 1985.

Sanger, F., Determination of nucleotide sequences in DNA. *Science*, 214, 1205-1210, 1981.

Singer, M., Highly repeated Sequences in mammalian genomes. *Int. Rev. Cytol.* 76, 67-112, 1982.

Singh, L., Purdom, I.F. and Jones, K.W., Effect of different denaturing agents on the detectability of specific DNA sequences of various base compositions by in situ hybridisation. *Chromosoma (Berl.)* 60, 377-389 1977.

Simmons, M.C., Maxwell, J., Haliotis, T., Higgins, M.J., Roder, J.C., White, B.N. and Holden, J.J.A., Amplified KpnI repetitive DNA sequences in homogeneously staining regions of a human melanoma cell line. *JNCI*, 72 (4), 801-808, 1984.

Smith, G.P., Evolution of repeated DNA sequences by unequal crossover. *Science* 191, 528-535, 1976.

Southern, E.M., Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.*, 98, 503-517, 1975.

Stark, G.R. and Wahl, G.M., Gene amplification. *Ann. Rev. Biochem.* 53, 447-491, 1984.

Stephenson, E.C., Erba, H.P. and Gall, J.G., Histone gene clusters of the newt *Notophthalmus* are separated by long tracts of satellite DNA. *Cell*, 24, 639-647, 1981.

Sumner, A.T., A simple technique for demonstrating centromeric heterochromatin. *Exp. Cell. Res.* 75, 304-306, 1972.

Temin, H.M., Reverse transcription in the eukaryotic genome: Retroviruses, pararetroviruses, retroposons, and retrotranscripts. *Mol.Biol.Evol.* 2 (6), 455-468, 1985.

Young, M.W. and Schwartz, H.E., *Cold Spring Harbour Symp. Quant. Biol.* 45, 629-40, 1981.

Conservation of highly repetitive DNA in cetaceans

Ü. Årnason, M. Höglund, and B. Widegren

Institute of Genetics, University of Lund, Sölvegatan 29, S-223 62 Lund, Sweden

Abstract. It is controversial whether odontocetes (toothed whales) and mysticetes (whalebone whales) have a common ancestry. Cetacean karyological uniformity, which is unique among mammalian orders, suggests a monophyletic origin; however, several anatomical authorities have maintained that odontocetes and mysticetes are diphyletic. We investigated the issue using Southern blot hybridization. Two labelled restriction fragment probes from the DNA of the sei whale (a mysticete) were hybridized to restricted DNA of cetacean species representing all extant families except the Eschrichtiidae, the gray whales. The probes hybridized to specific restriction fragments in all odontocete and mysticete materials. Hybridizations showed preservation of hybridization homologies and a striking conservation of the length of highly repeated DNA sequences. The results are compatible with a common ancestry of odontocetes and mysticetes.

Introduction

Cytogenetic investigations have shown marked karyological uniformity among cetaceans (Årnason 1974a; Duffield 1977). Most species have $2n=44$ chromosomes, and strikingly similar karyotypes. In the odontocetes (toothed whales), the Platanistidae (river dolphins), Delphinidae (dolphins, porpoises) and Monodontidae (narwhals) are characterized by the $2n=44$ karyotype. The Ziphiidae (beaked whales) and Physeteridae (sperm whales) have $2n=42$. The ziphiids share considerable similarities with the $2n=44$ karyotype, whereas the relationship between the $2n=44$ karyotypes and the physterids is less evident. Of the mysticetes (whalebone whales), the Eschrichtiidae (gray whales) and Balaenopteridae (rorquals) have $2n=44$. The Balaenidae (right whales) have a $2n=42$ karyotype that evolved from the $2n=44$ karyotype by fusion of two pairs of chromosomes (Jarrell 1979; Årnason et al. 1982).

Studies on banded and unbanded karyotypes have shown that in marine mammals karyological conservatism also occurs in the pinnipeds, i.e., the fur seals, sea lions, walrus, and true seals, (Fay et al. 1967; Årnason 1974b, 1981; Anbinder 1980). The evolutionary implications of this conservatism have been discussed previously (Årnason 1972, 1982a), the general conclusion being that a presumed low degree of inbreeding among marine mammals acts against the establishment of new karyological forms. Fac-

tors considered primarily responsible for karyological uniformity in marine mammals were: (1) low reproductivity due to late sexual maturity and not more than one offspring per year, (2) good vagility, and (3) an environment without distinct physical barriers. A relationship between karyotypic uniformity and size (Wilson et al. 1975) is not relevant in marine mammals, and no difference in the rate of karyotypic evolution in relation to size can be demonstrated. Thus, porpoises and balenopterids, which may differ in size by a factor of 10^3 , do not exhibit different rates of karyotypic evolution.

The karyological similarities among the cetaceans discussed above are particularly remarkable since anatomical and paleontological evidence cannot definitely prove whether the origin of Odontoceti and Mysticeti is monophyletic or diphyletic (Yablokov et al. 1972). The aim of the present study was to determine if cetacean karyological conformity was paralleled by preservation of highly repeated DNA common to both odontocetes and mysticetes. Homology at the DNA sequence level would constitute strong support for a common ancestry of these two groups.

Materials and methods

DNA preparation and analysis. DNA was prepared from frozen spleen or liver samples. For analytical purposes 2–3 µg DNA was digested with restriction endonucleases using conditions recommended by commercial suppliers (Boehringer Mannheim, BRL). When lengths of DNA fragments were determined, λ DNA digested with Hind III or Hind III + Eco RI was used as a marker. Digested DNAs were run on 1.0% or 1.5% agarose slab gels.

Physical mapping of cleavage sites. Locations of cleavage sites for Bcl I, Eco RV, Hind III, Pst I and Sac I in highly repetitive delphinid DNA were determined by digesting whole DNA with mixtures of two endonucleases. Lengths of fragments were determined using λ DNA digested with Hind III + Eco RI as a marker.

Preparation of labelled probes. A 200-µg sample of sei whale DNA was digested with Eco RV and run on a 1% Sea Plaque agarose gel (Marine Colloids). At termination of electrophoresis, the DNA in the gel was stained with ethidium bromide and two separate fragments were cut out under long-wave UV illumination. The lengths of the proximal and distal fragments were 1730 and 1430 base pairs

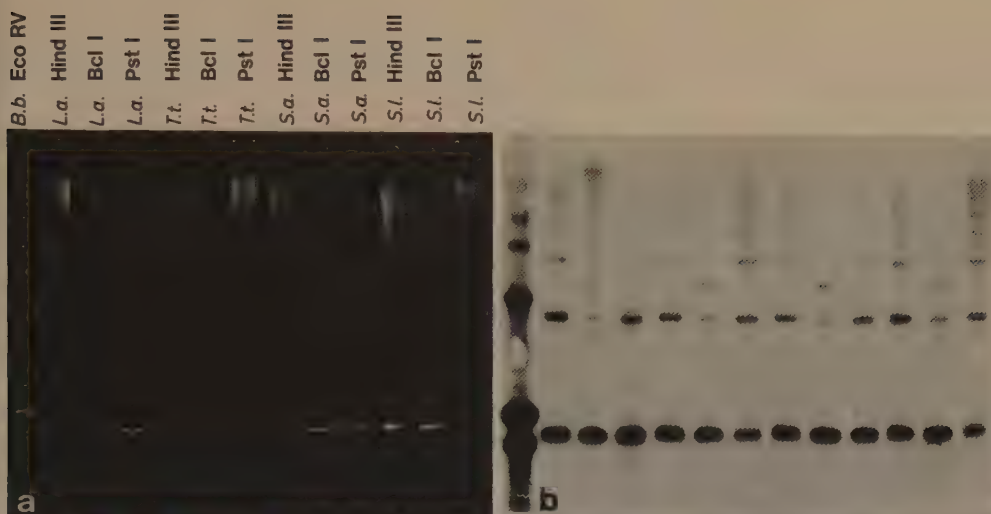


Fig. 1a, b. DNA of one mysticete (sei whale, *Balaenoptera borealis*, B.b.) and four delphinids (whitebeak dolphin, *Lagenorhynchus albirostris*, L.a.; bottlenose dolphin, *Tursiops truncatus*, T.t.; bridled dolphin, *Stenella attenuata*, S.a. and spinner dolphin, *S. longirostris*, S.l.). a After digestion with restriction endonucleases the fragments were separated by electrophoresis on 1% agarose gel. Digestion of sei whale DNA with Eco RV gave rise to two distinct restriction fragment bands with lengths of 1730 and 1430 bp. The delphinid DNAs were digested with Hind III, Bcl I, and Pst I. In all combinations a prominent fragment band, 1570 bp in length, was obtained. b The DNA in the gel was transferred to a nitrocellulose sheet according to the Southern blot technique. A ^{32}P nick-translated probe from the 1730-bp sei whale Eco RV fragment was hybridized to the transferred DNA. Hybridization was positive in the 1730- and 1430-bp sei whale fragments and multiples of the 1730-bp fragment. In the delphinid material hybridization was registered in the 1570-bp fragments and multiples thereof. In the Bcl I digests, hybridization to a specific fragment, about 3500 bp in length, was also recorded

(bp) respectively. After the agarose was melted, the samples were extracted with phenol and the DNA precipitated. 1 μg from each sample was nick translated with ^{32}P as described by Rigby et al. (1977).

DNA transfer and hybridization. Transfer of DNA from gels to nitrocellulose sheets (Millipore) was carried out according to Southern (1975). The filter sheets were prehybridized overnight at 65°C in $3 \times \text{SSC}$ ($1 \times \text{SSC} = 0.15 \text{ M}$ sodium chloride, 0.015 M sodium citrate), $10 \times$ Denhardt's solution ($1 \times$ Denhardt's solution = 0.02% ficoll, 0.02% polyvinylpyrrolidone, 0.02% bovine serum albumin). Hybridization was at 65°C in $3 \times \text{SSC}$, $1 \times$ Denhardt's solution. Washing was in $2 \times \text{SSC}$ at 65°C . Filters were exposed at room temperature to Kodak X-Omat RP X-ray film in Siemens cassettes with intensifying screens for 60–90 min.

Results

Restriction endonuclease digestion of total DNA of various delphinid species revealed the presence of single restriction sites for several restriction endonucleases within a highly repetitive unit 1570 bp in length. Figure 1a shows the results of digestion of DNA of four delphinids (*Lagenorhynchus albirostris*, *Tursiops truncatus*, *Stenella attenuata*; and *S. longirostris*). Three DNA samples of each species were digested with Hind III, Bcl I, and Pst I, respectively. In all cases, a conspicuous accumulation of 1570-bp fragments was observed. Identical fragment size was also obtained when Eco RV and Sac I were applied to the same materials

and to killer whale DNA. The DNA of one mysticete, the sei whale, *Balaenoptera borealis*, digested with Eco RV was also included in the electrophoresis. In this species as well as in other mysticete families (Balaenidae, Eschrichtiidae), digestion with Eco RV gave rise to two fragments, 1730 and 1430 bp in length. The 1730-bp sei whale fragment was ^{32}P nick translated and hybridized according to the Southern (1975) technique to the DNAs shown in Figure 1a. The results of the hybridization are shown in Figure 1b. The probe hybridized to the prominent 1730- and 1430-bp fragments of the sei whale and to multiples of the 1730-bp fragment. Hybridization to multiples of the 1430-bp fragment was not found. These results are consistent with two partially modified Eco RV sites in the repeated sequence with the distance of 300 and 1430 bp between sites. This localization of the Eco RV sites has been verified by studying 1730-bp fragments cloned from incomplete Eco RV digests of total DNA. The sei whale probe also hybridized in all four delphinid species to the 1570-bp fragment and multiples thereof. Hybridization was also recorded in a specific Bcl I fragment in the four delphinid DNAs; the size of this fragment was intermediate between the dimers and trimers of the 1570-bp fragment.

Figure 2a shows the results of digestion of DNA from 11 cetacean species by various restriction endonucleases, usually Eco RV. The materials represented all extant cetacean families except the gray whales. The samples designated 12c were the plasmid pBR322 with the 1570-bp fragment of the killer whale inserted in the Hind III site. All species except the two delphinids, the killer whale (*O. orca*)

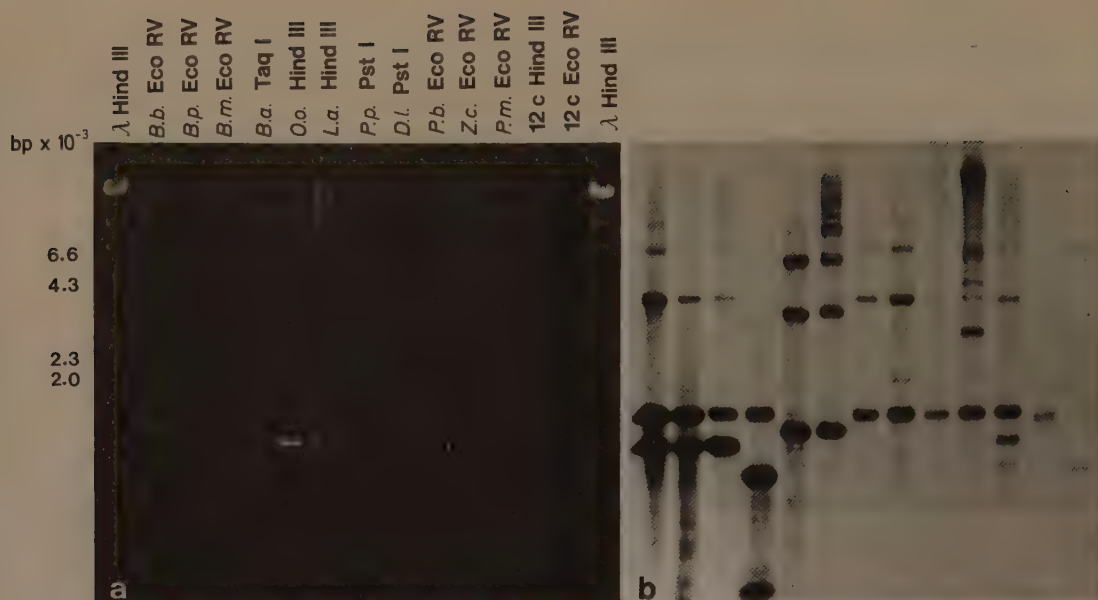


Fig. 2a, b. DNAs of 11 cetaceans, the 4 mysticete species (were sei whale, *Balaenoptera borealis*, B.b.; fin whale, *B. physalus*, B.p.; minke whale, *B. acutorostrata*, B.a. and the bowhead whale, *Balaena mysticetus*, B.m.). The 7 odontocete species (were killer whale, *Orcinus orca*, O.o.; whitebeak dolphin, *Lagenorhynchus albirostris*, L.a.; harbor porpoise, *Phocoena phocoena*, P.p.; belukha, *Delphinapterus leucas*, D.l.; La Plata river dolphin, *Pontoporia blainvillei*, P.b.; goosebeak whale, *Ziphius cavirostris*, Z.c., and sperm whale, *Physeter macrocephalus*, P.m.). A pBR322 recombinant (12c) with an inserted killer whale Hind III restriction fragment was also included. a After digestion with restriction endonucleases the fragments were separated by electrophoresis on 1% agarose gel. Each mysticete slot was loaded with 2.0–2.2 μ g DNA; each odontocete slot, with 3 μ g DNA, and each plasmid slot, with 0.1 μ g DNA. The initials of each species together with a designation of the restriction endonuclease used, are shown above each lane. A highly repetitive restriction fragment, 1730 bp in length, was observed in all families except the Delphinidae. In the delphinids (O.o. and L.a.) the 1730-bp fragment was replaced with a 1570-bp fragment. b Applying the Southern blot technique a 32 P nick-translated probe from the 1430-bp sei whale Eco RV fragment was hybridized to the transferred DNA. Hybridization was recorded in the 1730-bp fragment in all materials. Hybridization also occurred in the mysticete 1430-bp Eco RV fragments, the minke whale 1200-bp Taq I fragments, and the delphinid 1570-bp fragments

and the whitebeak dolphin (L.a.), showed restriction fragments 1730 bp in length.

In the hybridization following electrophoresis and Southern transfer a 32 P nick translated probe from the distal sei whale Eco RV fragment was used for hybridization. The results are shown in Figure 2b. The probe hybridized with both the proximal and distal Eco RV fragments in the two balenopterids (sei and fin whales). It also hybridized with corresponding fragments in the right whale, *Balaena mysticetus* (family Balaenidae). Hybridization also occurred in Taq I (1200 bp) fragments of the DNA of minke whale, *Balaenoptera acutorostrata*. In the odontocete samples, representing all extant families, distinct hybridization was registered to a 1730-bp repeat in all materials except the two species belonging to the family Delphinidae. In these two species, the killer whale (*O. orca*) and the whitebeak dolphin (*L. albirostris*), hybridization occurred in the prominent 1570-bp fragment. Barely recognizable hybridization to a 1730-bp fragment may, however, also have occurred in the two species. Hybridization was also recorded in the killer whale Hind III fragment of the clone 12c, but the cloned fragment was slightly larger than the 1570-bp repeat.

As mentioned above, single cleavage sites for several

restriction enzymes occur in delphinid highly repeated DNA. The relative locations of these sites were determined in four species after double digestion in various combinations. The results of some of these digestions are shown in Figure 3. The relative location of the sites was identical in all four species studied. The portion of 1570-bp fragments remaining after double digestion is presumed to be due to both heterogeneity between repeats and incomplete digestion. Digestion of various cloned fragments has revealed some heterogeneity in the repeated sequences. A physical map giving the location of the sites studied is shown in Figure 4.

Discussion

Previous molecular hybridizations in cetaceans showed the presence of related highly repetitive sequences in several species (Árnason et al. 1982; Árnason 1982b). Specific chromosomal localization of DNA satellites was also demonstrated (Árnason et al. 1978).

In the present study, restriction patterns of DNAs representing all cetacean families were investigated, and the hy-

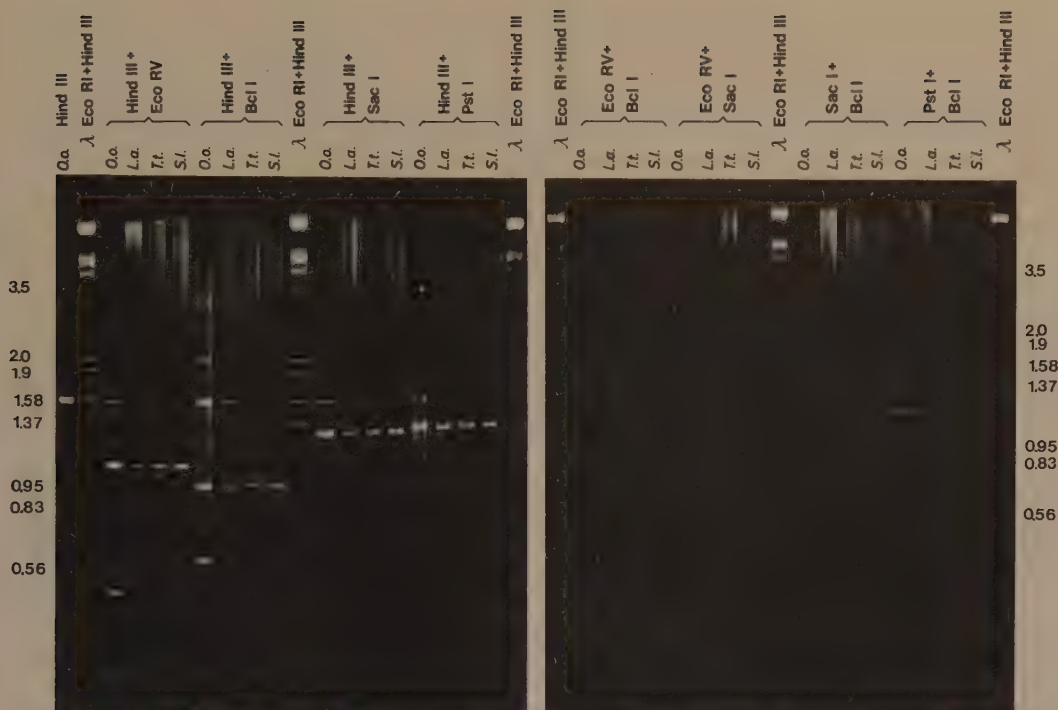


Fig. 3. DNA of four delphinids belonging to four different genera was digested with eight combinations of restriction endonucleases and run on agarose gel. Identical relative location of restriction sites was found in the highly repetitive 1570-bp fragment of all species. Designations of restriction endonucleases applied are shown above each set of digestions. *O.o.* killer whale; *L.a.* whitebeak dolphin; *T.t.* bottlenose dolphin; *S.l.* spinner dolphin. Marker fragments were obtained by digesting λ DNA with Eco RI + Hind III

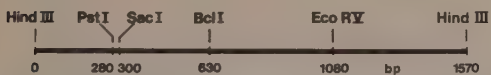


Fig. 4. Localization of restriction sites in delphinid 1570-bp highly repeated DNA. The numbers indicate the distance (bp) of different cleavage sites relative to the Hind III site

bridizations included representatives of all families except the Eschrichtiidae. A 1730-bp highly repetitive fragment characterized all cetacean families except the Delphinidae in which a 1570-bp fragment was found. In addition to the 1730-bp fragment digestion of mysticete DNA with Eco RV gave rise to a 1430-bp fragment. A highly repetitive fragment of this size was not found in any odontocete materials. The preservation of fragment length in cetacean highly repetitive DNA is striking when the evolutionary age of various families is considered. The relationship between the harbor porpoises, *Phocoena*, and the delphinids is not definitely settled, and several authors have argued against including *Phocoena* in the Delphinidae (Slipser 1936; Nishiwaki 1963; Mead 1975). The presence of a 1730-bp fragment in the harbor porpoise provides tentative support for this view.

The present hybridization studies have shown that besides maintenance of fragment length, substantial hybrid-

ization homologies have been preserved in highly repetitive DNA of all families during the evolution of the Cetacea. Identical locations of restriction sites in various delphinid DNAs also underline the apparently slow divergence of cetacean highly repetitive DNA.

Compared to other delphinids, a striking accumulation of heterochromatin has occurred in the killer whale (Arnason et al. 1980). The preservation of fragment length and relative locations of restriction sites in highly repetitive DNA of the killer whale show that this accumulation has taken place through amplification of the ancestral sequence without the manifestation of any major changes.

It is generally accepted (e.g., Lewin 1982) that there is a clear distinction between highly repetitive DNA of one species and that of another, no matter how closely related. The conservatism in the cetacean highly repetitive DNA found in this study does not seem to conform with this view.

As mentioned initially, it has not been possible on anatomical and paleontological grounds to definitely settle the question of whether odontocetes and mysticetes have a common ancestry. The present results show that related highly repetitive DNA sequences of the same length occur in both. Previous results from studies on cetacean karyology and molecular hybridization appear to be highly incompatible with a diphyletic origin of odontocetes and mysticetes.

At this stage it cannot be excluded that the conservatism in the divergence of cetacean highly repetitive DNA is due to a particular function of this class of DNA in the whale genome. However, if a common cause promotes conservation of highly repetitive DNA and karyotypes, it is probably the low degree of inbreeding that is thought to exist among marine mammals where the maintenance of small breeding isolates is unlikely.

Acknowledgements. We thank Kristján Loftsson, Hvalfjörður; Guðmundur Haraldsson, Akureyri; Agnar Thórisson, Hjalteyri; Helgi Hédinsson, Húsavík; Guðmundur Pétursson and Páll A. Pálsson, Keldur Laboratory, Iceland; Bengt O. Röken, Kolmården, Sweden; Grodon H. Jarrell, Fairbanks, Alaska; William F. Perrin and Kurt Benirschke, San Diego, California; Maria Cristina Pinedo, Rio Grande; and Hector Seuanez and Flavio Ortigao, Rio de Janeiro, Brazil, for materials and laboratory facilities. The work was supported by the Swedish Natural Science Research Council (Grant B-3655-103), the Tryggers Foundation, and the Erik Philip-Sörensen Foundation.

References

- Anbinder EM (1980) Karyology and evolution of pinnipeds, (in Russian) Nauka, Moscow, p 152
- Árnason Ú (1972) The role of chromosomal rearrangement in mammalian speciation with special reference to Cetacea and Pinnipedia. *Hereditas* 70:112-118
- Árnason Ú (1974a) Comparative chromosome studies in Cetacea. *Hereditas* 77:1-36
- Árnason Ú (1974b) Comparative chromosome studies in Pinnipedia. *Hereditas* 76:179-226
- Árnason Ú (1981) Localization of nucleolar organizing regions in pinniped karyotypes, *Hereditas* 94:29-34
- Árnason Ú (1982a) Karyotype stability in marine mammals. *Cytogenet Cell Genet* 33:274-276
- Árnason Ú (1982b) Southern blot hybridization in cetaceans, using killer whale restriction fragment as a probe. *Hereditas* 97:47-49
- Árnason Ú, Purdom IF, Jones KW (1978) Conservation and chromosomal localization of DNA satellites in balenopterid whales. *Chromosoma* 66:141-159
- Árnason Ú, Lutley R, Sandholt B (1980) Banding studies on six killer whales: An account of C-band polymorphism and G-band patterns. *Cytogenet Cell Genet* 28:71-78
- Árnason Ú, Purdom IF, Jones KW (1982) Cetacean molecular hybridization using balenopterid satellite DNA cRNA as probes. *Hereditas* 97:33-36
- Duffield DA (1977) Phylokaryotypic evaluation of the Cetacea. Ph.D. Thesis, Univ of California, Los Angeles
- Fay FH, Rausch VR, Feltz ET (1967) Cytogenetic comparisons of some pinnipeds (Mammalia: Eutheria). *Can J Zool* 45:773-778
- Jarrell GH (1979) Karyotype of the bowhead whale (*Balaena mysticetus*). *J Mammal* 60:607-610
- Lewin R (1982) Repeated DNA still in search of a function. *Science* 217:621-623
- Mead JG (1975) Anatomy of the external nasal passages and facial complex in the Delphinidae (Mammalia: Cetacea). *Smithson Contrib Zool* 207:1-72
- Nishiwaki M (1963) Taxonomical consideration on genera of Delphinidae. *Sci Rep Whales Res Inst* 17:93-104
- Rigby PW, Dieckmann M, Rhodes C, Berg P (1977) Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. *J Mol Biol* 113:236-251
- Slijper EJ (1936) Die Cetaceen, vergleichend-anatomisch und systematisch. *Capita Zool* 7:1-590
- Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* 98:503-517
- Wilson AC, Bush GL, Case SM, King M-C (1975) Social structuring of mammalian populations and rates of chromosomal evolution. *Proc Natl Acad Sci* 72:5061-5065
- Yablokov AV, Belkovich VM, Borisov VI (1972) Whales and dolphins (in Russian) Nauka, Moscow, p 472

Received August 1, 1983 / in revised form November 24, 1983

Accepted by W. Beermann

Studies on ribosomal DNA in cetaceans (whales)

ÚLFUR ÁRNASON and BENGT WIDEGREN

Institute of Genetics, University of Lund, Sweden

ÁRNASON Ú. and WIDEGREN, B. 1984. Studies on ribosomal DNA in cetaceans (whales).—*Hereditas* 101: 149–154. Lund, Sweden. ISSN 0018-0661. Received November 28, 1983

Ribosomal DNA (rDNA) of 4 mysticetes and 10 odontocetes were studied by Southern blot hybridization. The material represented all extant cetacean families except the gray whales. In all species, digestion with Eco RI gave rise to two distinct bands, 7.7–8.7 and 6.6–7.4 kb in size. The fin, sei and sperm whales were studied with respect to rDNA polymorphism. None of these species showed polymorphism in the distinct fragments. Outside the distinct fragments, bands of lower hybridization intensity were observed in the sei and sperm whales. In the sei whale, 3 out of 17 specimens differed from the general pattern. In the sperm whale, all 17 specimens studied showed different patterns of rDNA polymorphism in the faint bands.

Úlfur Árnason, *Institute of Genetics, University of Lund, Sölvegatan 29, S-223 62 Lund, Sweden*

Highly repetitive DNA of cetaceans exhibits striking conservatism with respect to both hybridization homologies and length of repeats (ÁRNASON et al. 1978, 1984; ÁRNASON 1982). Thus, substantial hybridization homologies occur between highly repetitive DNA of odontocetes (toothed whales) and mysticetes (baleen whales), two groups which separated some 45 million years ago. Identical lengths of repeats have also been preserved in both groups. This refers to a very highly repeated 1730 bp fragment which characterizes all cetacean families except the Delphinidae. The aim of the present study was to investigate whether conservatism of repeat lengths of DNA with lower repetitivity, namely ribosomal DNA, would differ from the conservation observed in very highly repetitive DNA.

In the fin whale the buoyant density of the 18S and 28S rDNA has been determined to 1.705 g/ml (ÁRNASON et al. 1977), but hybridization to rDNA digested with restriction endonucleases has not been performed previously.

The present material comprised 4 mysticete and 10 odontocete species. The mysticete species were the bowhead, minke, sei, and fin whales. The odontocete species were the La Plata river dolphin, the harbour porpoise, the belukha, the killer whale, the whitebeak dolphin, the bottlenose, spinner and bridled dolphins, the goosebeak and sperm whales. The sei, fin, and sperm whales were studied with respect to DNA polymorphism within each species. Of each species, 15–17 individuals were analysed. The analyses were primarily performed on DNAs digested with Eco RI.

Material and methods

DNA was extracted from solid tissues (liver, spleen), essentially according to the MARMUR (1961) procedure, with two pronase, RNase, and α -amylase (liver only) steps. DNA was digested with restriction endonucleases under conditions recommended by the manufacturer (Boehringer Mannheim). Digested DNA was run on 0.7% agarose gels submerged in $0.8 \times$ TEB. λ DNA digested with Hind III was used as a marker. Transfer of DNA to nitrocellulose filters (Schleicher and Schüll) was according to SOUTHERN (1975). Prior to transfer, upper parts of gels were exposed to UV-light at 254 nm. Plasmid pAT 153 with *Xenopus laevis* ribosomal cistron inserted in the Hind III site, was used for hybridization. The recombinant plasmid was a gift from Drs. E. Southern and A. Bird, Mammalian Genome Unit, University of Edinburgh. The plasmid was nick translated with ^{32}P dCTP to a specific activity of $5\text{--}8 \times 10^7$ cpm/ μg . Amount of nick translated probe used for each hybridization was 0.05–0.1 μg . Incubation was at 64°C in $3 \times$ SSC, $2 \times$ Denhardt's solution. Washing was in $1 \times$ SSC at 64°C. For autoradiography Kodak X-Omat X-ray film was used. The film was exposed in Siemens X-ray cassettes with intensifying screens for 1–3 days at -80°C .

Results

DNA of various species was digested with different restriction endonucleases, Bam HI, Eco RI, Eco

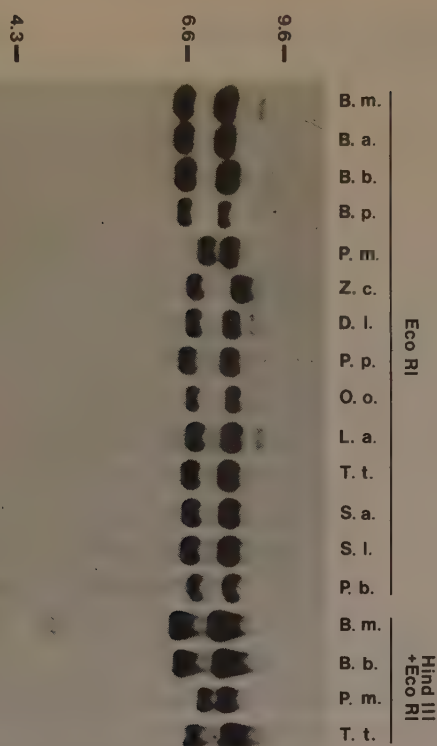


Fig. 1. DNA of 14 cetacean species was digested with Eco RI or Eco RI + Hind III. The DNA was run on a 0.7% agarose gel, transferred to a nitrocellulose membrane and hybridized to a ^{32}P labelled recombinant plasmid (pAT153+*Xenopus laevis* rDNA). B.m.=*Balaena mysticetus*, bowhead; B.a.=*Balaenoptera acutostrata*, minke whale; B.b.=*Balaenoptera borealis*, sei whale; B.p.=*Balaenoptera physalus*, fin whale; P.m.=*Physeter macrocephalus*, sperm whale; Z.c.=*Ziphius cavirostris*, goosbeak whale; D.l.=*Delphinapterus leucas*, belukha; P.p.=*Phocoena phocoena*, harbour porpoise; O.o.=*Orcinus orca*, killer whale; L.a.=*Lagenorhynchus albirostris*, whitebeak dolphin; T.t.=*Tursiops truncatus*, bottlenose dolphin; S.a.=*Stenella attenuata*, bridled dolphin; S.l.=*Stenella longirostris*, spinner dolphin; P.b.=*Pontoporia blainvillei*, La Plata dolphin.

λ DNA digested with Hind III was used as a marker. Lengths of fragments are in kb.

RV, Hind III, Pvu II, Sac I, Sal I and Sma I. On basis of the results of these digests, Eco RI was selected for the subsequent comparisons both between and within species.

Fig. 1 shows the results of a digestion of DNA of various cetacean species followed by hybridization with the labelled plasmid probe. The species included representatives of all extant cetacean families,

except the gray whales. The bowhead, minke, sei, and fin whales belong to the mysticetes (baleen whales), the other belong to the odontocetes (toothed whales).

After digestion with Eco RI, prominent hybridization occurred in two fragments in all species. In the mysticete species, the sizes of these fragments were determined to 7.7. and 6.6 kb respectively. In the odontocetes some differences in fragment sizes were observed between the various species studied. In the sperm whale, the sizes of the fragments were 8.2 and 7.4 kb respectively, in the beaked whale, the sizes were 8.7 and 7.0 kb. In the other odontocetes shown in Fig. 1, the size of the larger fragment measured approximately 8.4 kb in all species except the La Plata dolphin, in which the size was 8.6 kb. The smaller fragment measured 7.0 kb except in the harbour porpoise, in which it measured 6.8 kb, and in the La Plata dolphin, where the size was 7.2 kb.

A combined digestion with Eco RI + Hind III was carried out on DNA of four species, the bowhead, sei, and sperm whales, and the bottlenose dolphin. The sizes of the prominent fragments as seen after hybridization, were the same as after digestion with Eco RI only. No Hind III sites did, thus, occur within the boundaries of the prominent Eco RI fragments.

In some of the species studied, hybridization of low intensity was recorded outside the two prominent bands. In the bowhead, limited hybridization was observed corresponding to an 8.8 kb fragment size. In the minke whale, hybridization occurred at 9.5 and 4.4 kb. In the sei whale included, hybridization to two large fragments was observed; the size of the fragments corresponded to 9.0 and 9.4 kb respectively. In the sperm whale specimen, hybridization was recorded at two sites, with an approximate size of 4.6 and 4.9 kb respectively. In the belukha, the harbour porpoise, the killer whale, and the whitebeak dolphin, limited hybridization was recorded in each species in one fragment outside the prominent fragments. The size of this fragment varied between 8.6 and 9.0 kb among the species. In the La Plata river dolphin, hybridization was recorded to a fragment approximately 3.6 kb in length.

After the Eco RI + Hind III digest of the bowhead DNA, the faint Eco RI 8.8 kb band was replaced by hybridization to a fragment 4.8 kb in length. In the sei whale, hybridization to the 9.0 and 9.4 kb Eco RI fragments, was not observed after the Eco RI + Hind III digest; hybridization to two fragments, 4.9 and 5.3 kb in size, respectively, was, however, registered. In the sperm whale and the bottlenose dolphin, no difference was seen between

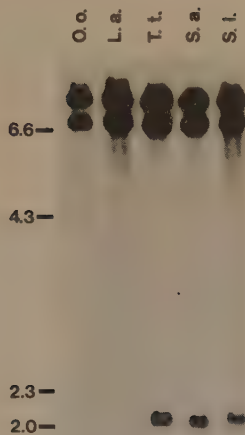


Fig. 2. Differences in rDNA of five delphinid species as observed after digestion with Eco RI. O.o.=killer whale; L.a.=whitebeak dolphin; T.t.=bottlenose dolphin; S.a.=bridled dolphin; S.l.=spinner dolphin. The 2kb fragment registered in T.t., S.a., and S.l. does not occur in L.a., and O.o.

the single Eco RI digest and the Eco RI + Hind III digest.

Hybridization to filters to which DNA fragments down to 1.5 kb length had been transferred, revealed some differences between the delphinid species studied. Thus, in the bottlenose, spinner and bridled dolphins, hybridization to a 2 kb long fragment was registered. Hybridization to a fragment of this size was not recorded in the killer whale and the whitebeak dolphin, Fig. 2.

As mentioned earlier, the recombinant plasmid used as a labelled probe, contained a complete *Xenopus laevis* rDNA unit. Due to the fact that probes containing either the 18S or the 28S unit were not available, we were unable to determine unequivocally the localization of Eco RI restriction sites within the coding part of the rDNAs studied. Conservation of the coding part of rDNA in different organisms has been documented in various studies (for references, see LONG and DAWID 1980) and similar localizations of restriction sites have been demonstrated, for example in *Xenopus laevis*, (WELLAUER et al. 1976) man (ERICKSON et al. 1981), rat (ROTHBLUM et al. 1982), and mouse (ARNHEIM and SOUTHERN 1977). In all these species, only two Eco RI sites occur within the coding region of the rDNA. One site is in the 3' end of the 18S region,

the other Eco RI site is in the 3' end of the 28S region. In man, the distance between these two sites has been determined to 7.3 kb (ERICKSON et al. 1981). The present results after cleaving various cetacean DNAs with Eco RI, are consistent with the presence of two Eco RI sites within the rDNA-coding region.

As described above, conspicuous hybridization occurred in two fragments in all species. Hybridization to the larger of these two fragments was somewhat more distinct than to the smaller fragment. On the basis of this we concluded, that the larger fragment contained a part of the cistron, extending from the 3' end of the 28S region into the 3' end of the 18S region, and that the smaller fragment extended from the 3' end of the 18S region, into the spacer region at the 5' end of the 18S region. The fragments showing a limited degree of hybridization are supposed to extend from the Eco RI site in the 28S region, into the spacer region at the 3' end of this unit. The diagram in Fig. 3 shows tentative localization of the Eco RI sites in the coding region and the adjacent spacer in rDNA of two cetacean species, the sei and sperm whales.

It is likely that, in cetaceans in general, the localization of the Eco RI sites in 18S and 28S is similar to those drawn in Fig. 3. This localization would give rise to hybridization at two distinct sites. In addition, low intensity hybridization would be expected in the fragment extending from the 3' end of the 28S region into the spacer region. In the non-coding part of this fragment, polymorphism might be expected. The fact that a fragment with a low degree of hybridization was not recorded in all species, may be explained by the presence of such a fragment being masked by the distinct fragments, or, that the size of the fragment was less than 1.5 kb and thus not a part of the DNA transferred to the nitrocellulose filters for hybridization.

In order to obtain information on the degree of rDNA polymorphism among the larger whales, a number of DNA samples of fin, sei and sperm whales were cleaved with Eco RI. Six selected specimens of each species are shown in Fig. 4. In the fin whale, no hybridization was observed outside the 7.7 and 6.6 kb fragments in the individuals included in Fig. 4 or in 9 additional specimens studied. In 17 sei whales studied, no variation was found in the size of the prominent 7.7 and 6.6 kb fragments. Outside these fragments, limited hybridization was registered in a 9.0 kb fragment in all individuals. Limited hybridization outside the 9.0 kb fragment was observed in three specimens. Two of these specimens had an additional fragment 8.2 kb in size,

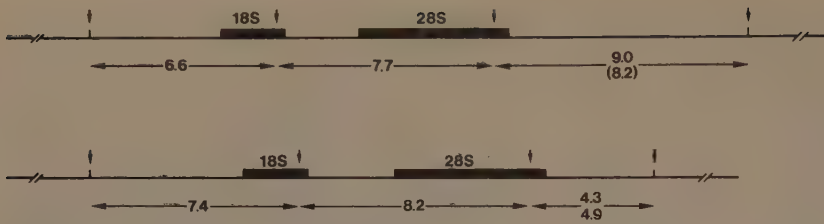


Fig. 3. Tentative localization of Eco RI sites in the 18S and 28S and adjacent regions in rDNA of the sei whale, above, and the sperm whale below.

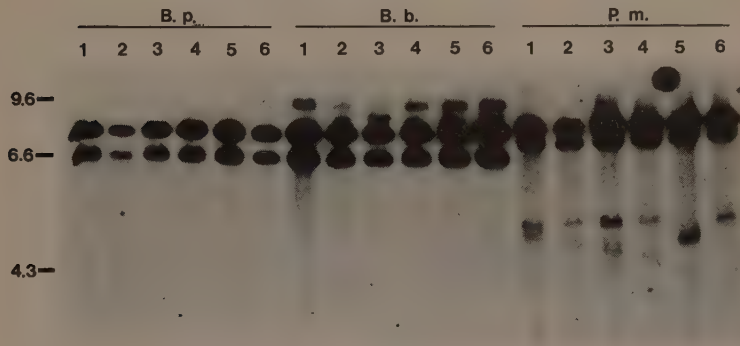


Fig. 4. DNA of six specimens of the fin, sei and sperm whales was digested with Eco RI. After hybridization no polymorphism was observed in the fin whale. In sei whales 1, 2 and 3, two faint bands were observed outside the two distinct Eco RI bands. In sei whale 1, the sizes of the faint bands were 9.0 kb and 9.4 kb, respectively. In sei whale 2 and 3, the corresponding sizes were 9.0 and 8.2 kb, respectively. In sei whales 4-6 showed the general pattern with one 9.0 kb band present. The sperm whale specimens all showed different patterns of polymorphism in the faint bands.

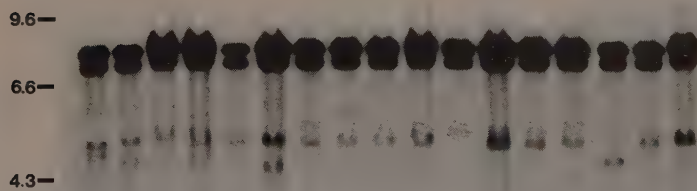


Fig. 5. DNAs of 17 male sperm whale specimens were digested with Eco RI. The bands with low hybridization intensity exhibit a high degree of polymorphism.

whereas in the third animal, the size of the additional fragment was 9.4 kb. These three specimens were included in Fig. 4 together with three specimens showing the general pattern. The sperm whale samples included in Fig. 4 all exhibited different polymorphisms with respect to sizes of fragments in

which limited hybridization was recorded. In the prominent 8.2 and 7.4 kb fragments no variation was observed.

The results of an extended study of a total number of 17 sperm whales are given in Fig. 5. A high degree of variation was observed in the sizes of fragments

which mostly fell into the 4–5 kb range. The number of fragments registered in each individual was usually 1–2, but a few specimens with 3 fragments were also found.

Discussion

Polymorphisms in ribosomal DNA have been described in various organisms, ex. *XENOPUS* (WELLAUER et al. 1976; BUONGIORNO-NARDELLI et al. 1977), mice (ARNHEIM and SOUTHERN 1977), humans (ARNHEIM and SOUTHERN 1977; WELLAUER and DAWID 1979; ERICKSON et al. 1981) and rats (ROTHBLUM et al. 1982). Studies of R-looped preparations of human rDNA (WELLAUER and DAWID 1979; ERICKSON et al. 1981) have shown that there is strong conservation of the transcribed rDNA, whereas the non-transcribed spacer rDNA may show considerable length heterogeneity.

The present study comprising four mysticete and ten odontocete species, showed, in general, similar sizes of the two prominent Eco RI rDNA fragments. Among the delphinids, a fragment approximately 2 kb in length was observed in the bottlenose, spinner and bridled dolphins. Hybridization to a fragment of this size was not observed in the killer whale and the whitebeak dolphin. The similarities between the bottlenose, spinner and bridled dolphins conform with the karyological identity of these species (ÁRNASON 1974, 1980) and anatomical characteristics (MEAD 1981).

The studies of variation within species was limited to three commercially caught species, the fin, sei, and sperm whales. In the fin whale, molecular hybridization did not show any rDNA polymorphism among the individuals studied. In the sei whale, three out of 17 individuals studied differed from a general pattern. In both these mysticete species, both sexes were represented. In the sperm whale, only male specimens were available. A high degree of rDNA polymorphism was observed among the sperm whales studied.

At this stage, there is no conclusive explanation for the difference between the two mysticete species and the sperm whale, with respect to the degree of rDNA polymorphism. It has been shown (ARNHEIM and SOUTHERN 1977), in man, that intragenomic rDNA polymorphisms reside at different chromosomal NOR regions. It is conceivable, that a higher number of NORs might contribute to a high degree of rDNA polymorphisms. In the fin and sei whales, there is only one NOR region per haploid genome

(ÁRNASON 1981) and it might be argued that this condition may contribute to the relative rDNA uniformity of these species. However, the sperm whale also has only the one NOR region per haploid genome, so this does not seem to be a valid point with respect to the three species studied. The NOR-carrying chromosomes of the sperm whale and the mysticetes are radically different. In the sperm whale this chromosome is very large, whereas in the two mysticetes in question it is the smallest chromosome of the complement. Between the NOR homologues of the mysticetes, associations are frequently observed; in the sperm whale associations between the NOR chromosomes are extremely rare. It is possible that in the mysticetes meiotic exchanges between NOR regions may contribute to homogenization of rDNA as compared with the sperm whale.

As mentioned earlier, the number of fragments in the 4–5 kb range in the sperm whale DNA varied between 1–3. As there are only two NOR regions per diploid cells in this species, it is obvious that each NOR region cannot be homogeneous with respect to the size of the rDNA unit.

With respect to population structure, there is a marked difference between the three species in question, as sperm whales form schools with single sexually mature males. School formations of this kind do not occur among the mysticetes. Whether this factor could have an effect on the different degree of rDNA conservation observed cannot, however, be postulated.

As mentioned in the introduction, a repetitive 1730 bp fragment has been found in all cetaceans except the delphinids, where the length of the repeat is 1570 bp (ÁRNASON 1982; ÁRNASON et al. 1984). In this highly repetitive fragment, hybridization homologues and restriction sites have been strongly conserved. This fragment may total 10–15% of the genome. The present study shows that, in general, there is a limited variation in cetacean rDNA. It is, however, noteworthy that the rDNA shows a greater degree of variation than the very highly repetitive 1730 bp fragment occurring in the same species.

Acknowledgments. — Most of the materials studied were collected at the whaling station in Hvalfjörður, Iceland. We thank the director Kristján Loftsson, Björn P. Þórhallsson, Þórdur B. Gudjónsson, Magnús D. Ólafsson and Andrés Magnússon for the generous collection of the various fin, sei and sperm whale samples. The recombinant plasmid (pAT 153+*X. laevis* ribosomal cistron) used was a gift from Drs. E. Southern and A. Bird, MRC, Mammalian Genome Unit, Edinburgh. The work was supported by the Swedish Natural Science Research Council (Grants B-3655-194 and -105) and the Erik Philip-Sörensen Foundation.

Literature cited

- ÁRNASON, Ú. 1974. Comparative chromosome studies in Cetacea. — *Hereditas* 77: 1–36
- ÁRNASON, Ú. 1980. C- and G-banded karyotypes of three delphinids: *Stenella clymene*, *Lagenorhynchus albirostris* and *Phocoena phocoena*. — *Hereditas* 92: 179–187
- ÁRNASON, Ú. 1981. Localization of NORs in cetacean karyotypes. — *Hereditas* 95: 269–275
- ÁRNASON, Ú. 1982. Southern blot hybridization in cetaceans, using killer whale restriction fragment as a probe. — *Hereditas* 97: 47–49
- ÁRNASON, Ú., LIMA-DE-FARIA, A., GRANSTRÖM, H. and ISAKSSON, M. 1977. Analysis of cetacean DNA II. Localization of 18S and 28S ribosomal RNA cistrons in a heavy DNA component. — *Hereditas* 87: 67–76
- ÁRNASON, Ú., PURDOM, I. F. and JONES, K. W. 1978. Conservation and chromosomal localization of DNA satellites in balenopterid whales. — *Chromosoma* 66: 141–159
- ÁRNASON, Ú., HÖGLUND, M. and WIDEGREN, B. 1984. Conservation of highly repetitive DNA in cetaceans. — *Chromosoma* 89: 238–242
- ARNHEIM, N. and SOUTHERN, E. M. 1977. Heterogeneity of the ribosomal genes in mice and men. — *Cell* 11: 363–370
- BUONGIORNO-NARDELLI, M., AMALDI, F., BECCARI, E. and JUNAKOVIC, N. 1977. Size of ribosomal DNA repeating units in *Xenopus laevis*: Limited individual heterogeneity and extensive population polymorphism. — *J. Mol. Biol.* 110: 105–117
- ERICKSON, J. M., RUSHFORD, C. L., DORNEY, D. J., WILSON, G. N. and SCHMICKEL, R. D. 1981. Structure and variation of human ribosomal DNA: molecular analysis of cloned fragments. — *Gene* 16: 1–9
- LONG, E. O. and DAWID, I. B. 1980. Repeated genes in eucaryotes. — *Annu. Rev. Biochem.* 49: 727–764
- MARMUR, J. 1961. A procedure for the isolation of deoxyribonucleic acid from microorganisms. — *J. Mol. Biol.* 3: 208–218
- MEAD, J. G. 1981. Anatomy of the external nasal passages and facial complex in the Delphinidae (Mammalia: Cetacea). *Smithsonian Contrib. Zool.* 207: 1–72
- ROTHBLUM, L. I., PARKER, D. L. and CASSIDY, B. 1982. Isolation and characterization of rat ribosomal DNA clones. — *Gene* 16: 75–77
- SOUTHERN, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. — *J. Mol. Biol.* 98: 503–507
- WELLAUER, P. K., DAWID, I. B., BROWN, D. D. and REEDER, R. H. 1976. The molecular basis for length heterogeneity in ribosomal DNA from *Xenopus laevis*. — *J. Mol. Biol.* 105: 461–486
- WELLAUER, P. K. and DAWID, I. B. 1979. Isolation and sequence organization of human ribosomal DNA. — *J. Mol. Biol.* 128: 289–303

Different rates of divergence in highly repetitive DNA of cetaceans

ÚLFUR ÁRNASON and BENGT WIDEGREN

Institute of Genetics, University of Lund, Sweden

ÁRNASON, Ú. and WIDEGREN, B. 1984. Different rates of divergence in highly repetitive DNA of cetaceans. — *Hereditas* 101: 171–177. Lund, Sweden. ISSN 0018-0661. Received December 2, 1983

Three highly repetitive components of cetacean genomes were studied. Two of the components were isolated from balenopterid DNA by isopycnic ultracentrifugation. These components were referred to as H (heavy) and L (light) satellites, respectively. The third component was isolated as a restriction fragment, 1730 bp in length. Molecular hybridization and restriction analysis showed that the 1730 bp fragment was highly conserved in both mysticetes and odontocetes, two lineages that separated some 45 million years ago. Digestion with Xba I showed that the repeat length of the H satellite was 420 bp in the three balenopterids studied. Sequences related to the H satellites occur outside the balenopterids. The L satellite is a balenopterid characteristic. The L satellite did not hybridize with DNA of species outside the genus *Balaenoptera*. Irrespective of the supposed younger age of the L satellite, hybridization with this component to restricted DNA showed great variation between the three balenopterid species. The L satellite is evidently evolving much faster than the other two components.

Úlfur Árnason, *Institute of Genetics, University of Lund, Sölvegatan 29, S-223 62 Lund, Sweden*

Constitutive heterochromatin makes up a considerable part of the balenopterid karyotypes as made visible by C-band technique. As much as 25–30% of the total length of the karyotypes are C-band positive (ÁRNASON 1974). Two highly repetitive components have been isolated by isopycnic centrifugation and their hybridization homologies studied (ÁRNASON et al. 1978). More recently, it was shown that a large portion of the balenopterid genome was made up by highly repetitive DNA with a repeat length of 1730 bp (ÁRNASON et al. 1984). In the present study, homologies within each highly repetitive family were studied by applying SOUTHERN blot (1975) technique. The three balenopterid species included were the minke whale, *Balaenoptera acutorostrata*, the sei whale, *B. borealis* and the fin whale, *B. physalus*. In the 1730 bp fragment, relative localization of several single restriction sites was also determined. The localizations of some of these sites were identical in both mysticetes and odontocetes.

Material and methods

DNA was prepared from spleen tissues, according to MARMUR (1961), including two RNase and pronase treatments.

Silver-caesium sulphate gradients were prepared

according to JENSEN and DAVIDSON (1966). Satellite DNAs were isolated by successive Ag/Cs₂SO₄ (Ag⁺ to DNA-P ratio 0.20 or 0.25) centrifugations in the M.S.E. 8 × 40 Ti rotor loading 700–1000 µg DNA per tube.

Recombinant plasmid 2106 was obtained by inserting a highly repetitive 1730 bp fragment of the bowhead into the Eco RV site of pBR322.

Satellite DNAs and recombinant plasmid 2106 were nick translated to high specific activity according to RIGBY et al. (1977).

Agarose gel concentrations were 1%. Transfer of DNA to nitrocellular filter membranes (Schleicher and Schüll) was according to SOUTHERN (1975). Hybridization was at 64°C in 3 × SSC, 2 × DENHARDT'S (1966) solution. Washing was in 1 × SSC at incubation temperature. Exposure of autoradiographic films was at room temperature in Siemens cassettes with intensifying screens.

Results

Minke whale DNA was digested with several restriction endonucleases (Eco RV, Eco RI, Cfo I, Taq I, Hae III, Alu I and Rsa I) and run on a 1% agarose gel. After transfer of the DNA (SOUTHERN 1975), the nitrocellulose membrane was used for subsequent hybridizations with the recombinant



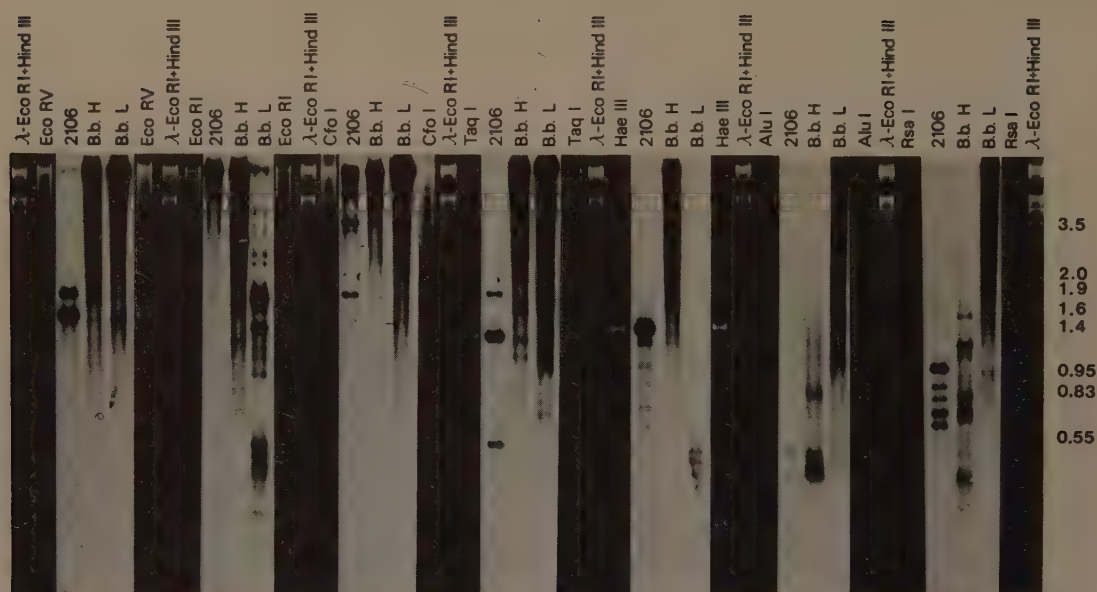


Fig 1. DNA of the minke whale, *Balaenoptera acutostrata* was digested with different restriction endonucleases and run on a 1% agarose gel. After Southern transfer of the DNA, three separate hybridizations to the filter were performed and the results compiled. For each digest the pattern after electrophoresis is given together with the results of the hybridization with the recombinant plasmid 2106 and the H and L satellites respectively. λ DNA digested with Eco RI + Hind III was used as a marker DNA; size of marker fragments is in kb.

plasmid 2106 and each of two balenopterid satellites. The satellites are referred to as B.b.H (sei whale heavy satellite, $q=1.710/1.711$) and B.b.L (sei whale light satellite $q=1.702/1.703$). The results of the hybridization are shown in Fig. 1. The figure is composed of lanes with marker DNA (λ digested with Eco RI + Hind III), minke whale DNA digested with the restriction endonucleases indicated, and the autoradiographic results of the three separate hybridizations. Between hybridizations the filter was washed in $0.5 \times \text{SSC}$, 0.5% SDS at 80°C for 30 min in order to release from the filter the labelled probe from the previous hybridization. The fragments with which the 2106 recombinant plasmid hybridized, were in all cases clearly visible in the EtBr stained gels. The Eco RI bands to which the L satellite hybridized were also visible on gels, but no accumulation of fragments was detected at the hybridization sites of the H satellite except when Alu I was used.

The highly repetitive 1730 bp fragment of the bowhead which was used for constructing the recombinant plasmid hybridizes with fragments of identical length found among both mysticetes and odontocetes. The occurrence of this fragment

among cetaceans has been discussed previously (ÁRNASON 1982a; ÁRNASON et al. 1984).

In order to determine the relative localization of restriction sites in the 1730 bp fragment among the balenopterids, series of double digests were performed. Some of the results are given in Fig. 2, in which the striking conformity between the three species is apparent. The map in Fig. 3 is based on the results in Fig. 2 plus additional digestions which are not shown.

Fig. 4 gives the results of some double digests of DNA of one balenopterid, the sei whale and two odontocetes, the harbour porpoise and the belukha. The two odontocetes belong to different odontocete families. In the 1730 bp highly repetitive fragment of the three species identical relative distances between restriction sites were found.

In order to obtain information on the conservation of the DNA to which the H and L satellites hybridized, a series of Southern blot hybridizations were undertaken in which labelled H and L satellite DNAs were used as probes. Fig. 5 gives the results of a hybridization of the H satellite to DNA digested with some different restriction endonucleases. DNA from the bowhead and the minke, sei, and fin

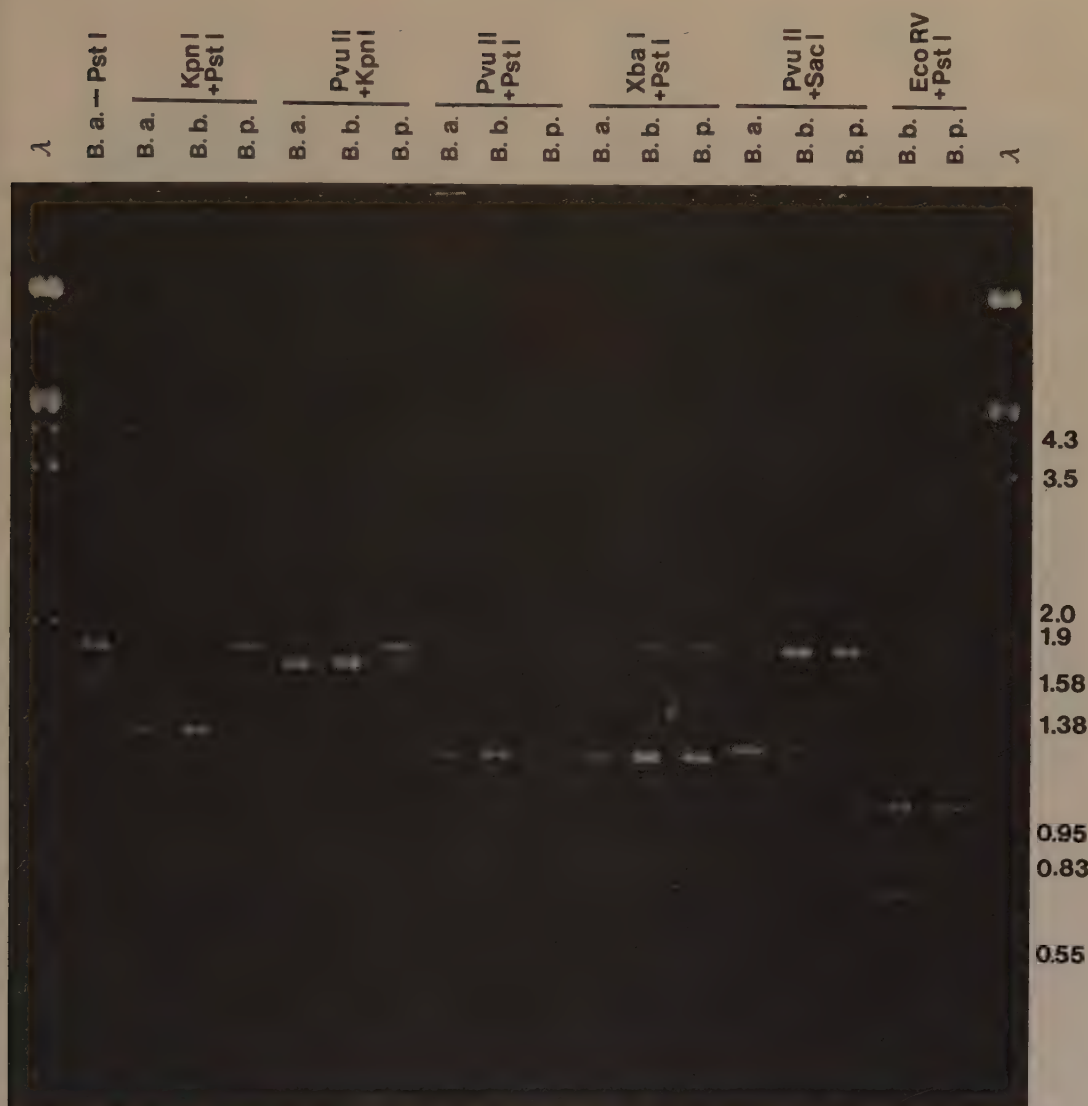


Fig 2. Results of digests of DNA of three balenopterid species. B.a. = *Balaenoptera acutorostrata*, minke whale; B.b. = *B. borealis*, sei whale; B.p. = *B. physalus*, fin whale. The 1730 bp fragment is conspicuous in the Pst I digest of the B.a. DNA. The double digest showed identical distances between restriction sites in the 1730 bp fragment of the three species.

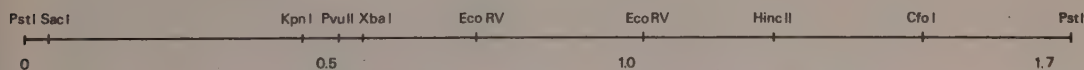


Fig. 3. Restriction map showing localization of 8 single restriction sites in the highly repetitive 1730 bp balenopterid unit.

whales were included in the Eco RI digestions. Other digestions were on fin whale DNA. The labelled sei whale H satellite hybridized to an Xba I repeat unit with a monomeric length of approxima-

tely 420 bp. Limited hybridization was observed in the Eco RI digests of the balenopterids. The ladder patterns in the three species were virtually identical. Some hybridization to the bowhead DNA did

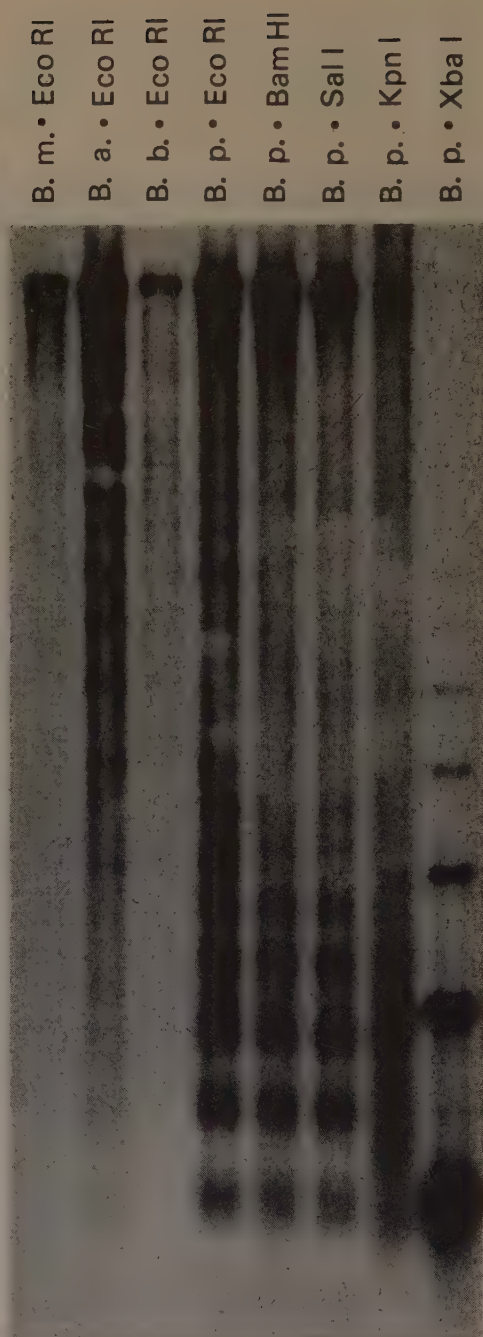


Fig. 4. Results of restriction digestion of DNA of one baleen whale and two toothed whales. B. b. = *Balaenoptera borealis*, sei whale; P. p. = *Phocoena phocoena*, harbour porpoise; D. l. = *Delphinapterus leucas*, belukha. The 1730 bp fragment is conspicuous in the Pst I digests. The double digests showed identical relative localization of restriction sites in the 1730 bp fragment in the three species.

Fig. 5. Autoradiographic results after hybridizing the heavy balenopterid satellite to baleen whale DNAs. B. m. = *Balaena mysticetus*, bowhead. Other abbreviations of species are as in Fig. 1. The length of the Xba I monomer unit is about 420 bp. In B. a., B. b. and B. p. partial digestion was obtained with Eco RI, the ladder pattern in the three species was the same. The heavy satellite hybridizes to B. m. DNA but no Eco RI ladder pattern occurs in this species.

occur, but no ladder pattern was recognized in this species.

Hybridization to Xba I digested DNA of the three balenopterids did not reveal any differences between



the sizes of the monomeric unit in the three species, Fig. 6. No hybridization was registered in the 1730 bp repeat.

Hybridization with the labelled L sei whale satel-

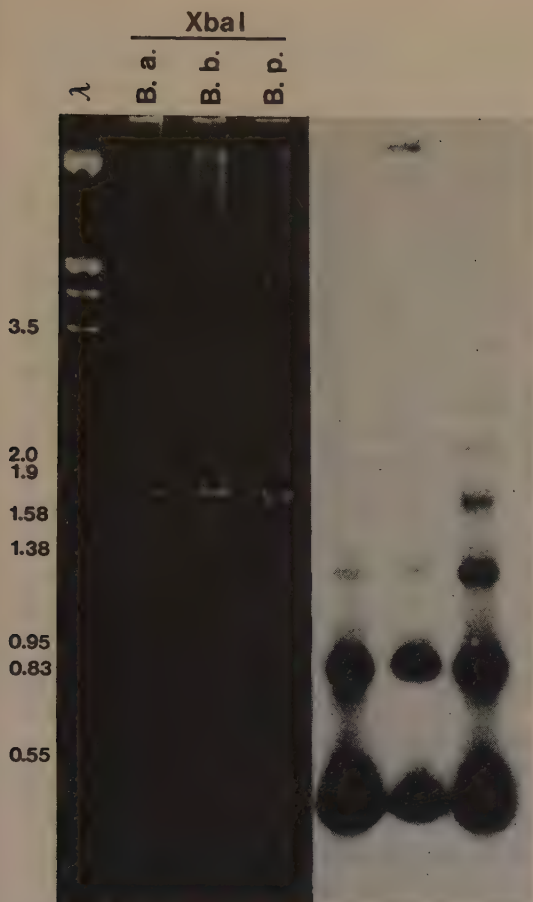


Fig. 6. Digestion of balenopterid DNA with Xba I gives rise to highly repetitive 1730 bp fragment; a repeat about 420 bp in length is also present but comprises a much less portion of the genome. The heavy satellite hybridizes with the 420 bp unit and multiples of this repeat. The satellite does not hybridize with the 1730 bp fragment.

lite revealed striking differences between the Eco RI digests of the three balenopterid species. Hybridization to other digests was also recorded, notably Xba I, Bcl I and Bgl II. Hybridization to the bowhead DNA was nonexistent, Fig. 7.

As in the Eco RI digests, hybridization with the L satellite to DNA digested with other restriction endonucleases also revealed distinct differences between the three balenopterids. Thus, conspicuous differences between hybridizations to Bcl I, Bgl II, Sau 3A and Mbo I digested DNA of the three species were observed. In each species hybridiza-

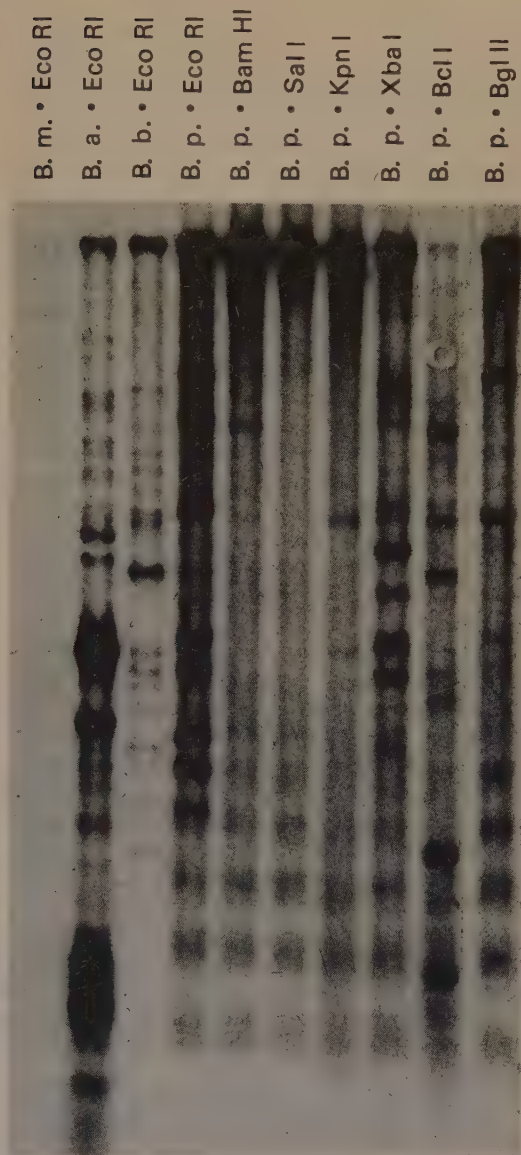


Fig. 7. Autoradiographic results after hybridizing the light balenopterid satellite to baleen whale DNAs. Hybridization to bowhead, B.m., DNA is nonexistent. In the Eco RI digests the difference between the three balenopterid species is striking.

tions to the Sau 3A and Mbo I digests were identical. This shows that the differences between the three species are not caused by differences in methylation of adenosine, which does affect digestion with Mbo I but not with Sau 3A.

Discussion

The H (heavy) and L (light) balenopterid DNA satellites have been studied in some detail, both with respect to hybridization homologies and to chromosomal distribution (ÁRNASON et al. 1978). The analysis did not reveal any hybridization homology between the two satellites, but within each satellite group, melting profiles were very similar, particularly in the case of the H satellite. In situ hybridization showed that the H satellite was predominantly located in terminal C-bands throughout the karyotypes, whereas the L satellite was found in centromeric C-bands in just a few pairs. The L satellite characterizes the balenopterids and no hybridization to odontocete DNA has been recorded. The H satellite hybridizes to DNA of cetaceans outside the balenopterids and in situ hybridization showed that, in the family Balaenidae, the general localization of the satellite was the same as in the balenopterids, i.e. in terminal C-bands (ÁRNASON et al. 1982).

The 1730 bp repeat is a cetacean characteristic which occurs in all cetacean families except the Delphinidae, in which the repeat length of a related component is 1580 bp (ÁRNASON et al. 1984). The repeat, which contains single restriction sites for several restriction endonucleases, comprises a large part of the genomes, and is conspicuous after electrophoresis of total DNA digested with several restriction endonucleases which have single or few restriction sites in the repeat.

The components to which the H and L satellites hybridize comprise a much less portion of the genome than the 1730 bp repeat. However, both can be made visible on moderately loaded gels as an Xba I repeat to which H hybridizes, and as an Eco RI repeat to which L hybridizes.

As mentioned above, filter hybridization between the L balenopterid satellite and odontocete DNA did not reveal the presence of related sequences among the odontocetes. In the case of the H satellite, varying degrees of hybridization between odontocete DNAs and this satellite have, however, been recorded (ÁRNASON et al. 1982). In the present Southern blot hybridizations, a very limited degree of hybridization was registered between the L satellite DNA and the bowhead DNA. The L satellite is therefore considered to be a balenopterid characteristic of a considerably younger age than the 1730 highly repetitive fragment, which occurs among both odontocetes and mysticetes.

There have been different opinions on the cetacean phylogeny (for references see YABLOKOV et al.

1972) but the similar karyology of odontocetes and mysticetes (ÁRNASON 1969, 1974; DUFFIELD 1977) together with the presence of common highly repetitive sequences, speaks definitely in favour of a monophyletic origin of cetaceans. The cetacean phylogenetic record is somewhat incomplete, but accepting a monophyletic origin of cetaceans, the divergence into odontocetes and mysticetes would probably have taken place some 45 million years ago. Irrespective of the exact date of the cetacean divergence, it is obvious that the age and conservation of the 1730 cetacean repeat is most respectable.

According to YABLOKOV et al. (1972) the family Balaenopteridae, to which the genus *Balaenoptera* belongs, can be traced back about 15 million years. The age of genus *Balaenoptera* is less, probably 7–9 million years.

The balenopterid data presented here are particularly interesting in the light of the strikingly different rates of divergence in the balenopterid L satellite when compared with the 1730 bp sequence. Thus, the L satellite shows a much greater diversity between species than the much older 1730 bp repeat.

The study of highly repetitive DNA in the marine mammals yields a particular opportunity to study the evolution of repetitive sequences among animals where the occurrence of small isolates is virtually nonexistent. The karyological conservation in the marine mammals has been attributed to this fact (ÁRNASON 1972, 1982b), and the cetacean system may prove to be a valuable source for investigating the issue of concerted evolution and molecular drive as discussed particularly by DOVER in a series of papers (e.g., DOVER et al. 1982).

Acknowledgements. — The work was supported by the Swedish Natural Science Research Council (Grants B-3655-104 and -105) and the Nilsson-Ehle Fund of the Royal Physiographic Society, Lund.

Literature cited

- ÁRNASON Ú. 1969. The karyotype of the fin whale. — *Hereditas* 62: 223–284
- ÁRNASON, Ú. 1972. The role of chromosomal rearrangement in mammalian speciation with special reference to Cetacea and Pinnipedia. — *Hereditas* 70: 113–118
- ÁRNASON, Ú. 1974. Comparative chromosome studies in Cetacea. — *Hereditas* 77: 1–36
- ÁRNASON, Ú. 1982a. Southern blot hybridization in cetaceans using killer whale restriction fragment as a probe. — *Hereditas* 97: 47–49
- ÁRNASON, Ú. 1982b. Karyotype stability in marine mammals. — *Cytogenet. Cell Genet.* 33: 274–276
- ÁRNASON, Ú., PURDOM, I.F. and JONES, K.W. 1978. Conserva-

- tion and chromosomal localization of DNA satellites in balenopterid whales. — *Chromosoma* 66: 141–159
- ÁRNASON, Ú., PURDOM, I.F. and JONES, K.W. 1982. Cetacean molecular hybridization using balenopterid satellite DNA cRNAs as probes. — *Hereditas* 97: 33–36
- ÁRNASON, Ú., HÖGLUND, M. and WIDEGREN, B. 1984. Conservation of highly repetitive DNA in cetaceans. — *Chromosoma* 89: 238–242
- DENHARDT, D.T. 1966. A membrane filter technique for the detection of complementary DNA. — *Biophys. Res. Commun.* 23: 641–646
- DOVER, G., BROWN, S., COEN, E., DALLAS, J., STRACHAN, T. and TRACK, M. 1982. The dynamics of genome evolution and species differentiation. — In *Genome Evolution, The Systematics Association Special Volume No. 20*, (Eds. G. DOVER and R. FLAVELL), Academic Press
- DUFFIELD, D. A. 1977. Phylokaryotypic evaluation of the Cetacea. — *Ph. D Thesis, Univ. of California, Los Angeles*
- JENSEN, R. H. and DAVIDSON, N. 1966. Spectrophotometric, potentiometric and density gradient ultracentrifugation studies of the binding of silver ion by DNA. — *Biopolymers* 4: 17–32
- MARMUR, J. 1961. A procedure for the isolation of deoxyribonucleic acid from microorganisms. — *J. Mol. Biol.* 3: 208–218
- RIGBY, P. W. J., DIECKMANN, M., RHODES, C. and BERG, P. 1977. Labelling of deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. — *J. Mol. Biol.* 113: 237–251
- SOUTHERN, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. — *J. Mol. Biol.* 98: 503–517
- YABLOKOV, A. V., BELKOVICH, V. M. and BORISOV, V. I. 1972. Whales and dolphins. — “Nauka”, Moscow. (In Russian)

Characteristics of a Conserved 1,579-bp Highly Repetitive Component in the Killer Whale, *Orcinus orca*¹

Bengt Widegren,* Úlfur Árnason,* and Göran Akusjärvi†

*University of Lund, Sweden; and †University of Uppsala, Sweden

A tandemly organized, highly repetitive DNA component of the killer whale was sequenced. The length of the repeat was 1,579 bp. This unit, which characterizes all delphinids, shows stringent hybridization homology with a 1,740-bp repeat that is characteristic of all other cetacean families. The 1,579-bp component comprises ~15% of the killer-whale genome, in which it is repeated $4-5 \times 10^5$ times. Computer analysis of the sequence showed no linear repetition within the component. This indicates that the 1,579-bp unit has not evolved by amplification of shorter repeats. Several inverted repeats of substantial length were found in the 1,579-bp unit. The most conspicuous of these was a 72-bp sequence that deviated from matching in only three positions. The 72-bp sequence occurs within an open reading frame 330 bp in length. Transcriptional activity was registered in the cloned repeat in a cell-free system. The length of the transcript was ~340 nucleotides. The chromosomal localization of the 1,579-bp repeat was determined by in situ hybridization. The repeat was present in eight of 21 autosomal pairs and was found in almost all C-band-positive (constitutive heterochromatin) regions of the karyotype.

Introduction

An evolutionarily conservative, highly repetitive component is present in all cetaceans (whales, dolphins, and porpoises). The odontocete (toothed whale) and mysticete (whalebone whale) lineages separated ~40–50 Myr ago, but it is conceivable that the component is more ancient than the separation of the odontocete and mysticete lineages. The length of the repeat is ~1,740 bp in all cetacean families except the Delphinidae, in which the length of the repeat is ~1,570 bp (Árnason et al. 1984). Hybridization between the 1,740-bp repeat and the delphinid repeat occurs under stringent conditions. It is logical to assume that the 1,740-bp repeat was replaced by the delphinid repeat prior to the radiation of the delphinids and that the age of the repeat thus exceeds 20 Myr.

As judged on the basis of restriction analysis, the length of the delphinid repeat and the relative localization of restriction sites have been preserved among the various delphinid genera. These characteristics of the repeat have thus been maintained despite different degrees of amplification of the repeat in different delphinid genera.

The repeat is particularly abundant in the killer whale, and this species also has a C-band pattern that distinguishes it from the pattern in other delphinid karyotypes. Thus, in the killer whale, C-bands are more prominent than in other delphinids and occur in chromosome positions unique to this species.

The present study was undertaken to characterize the highly repetitive delphinid

1. Key words: highly repetitive DNA, base composition, cetacean evolution.

Address for correspondence and reprints: Dr. Úlfur Árnason, Department of Genetics, University of Lund, Sölvegatan 29, S-223 62 Lund, Sweden.

Mol. Biol. Evol. 2(5):411–419. 1985.

© 1985 by The University of Chicago. All rights reserved.
0737-4038/85/0205-0355\$02.00

component as represented by the killer whale, with respect to both its composition and its chromosomal localization. The in situ hybridization technique was applied to clarifying whether the delphinid repeat was present in C-bands in positions unique for the killer whale as well as in positions common to other delphinids.

Experimental Procedures

Single sites for several restriction enzymes suitable for cloning occur in the presently studied delphinid repeat (Árnason et al. 1984). Cloning was performed after cleaving total DNA of the killer whale with *EcoRV*. The DNA was loaded on 1% agarose gel, and the highly repetitive fragment was electroeluted. After purification, the fragment was ligated into the *EcoRV* site of pBR322. Both strands of the cloned fragment were sequenced following the Maxam and Gilbert (1980) procedure after 3' and 5' end labeling. Positions 1 and 1,579, respectively, refer to axis of the *EcoRV* site in the repeat.

Computer analysis of the repeat was performed using three different programs. One program was provided by Drs. Petter Gustafsson and Robert Harr (Dept. of Microbiology, Umeå, Sweden), another by Dr. Philip Taylor (MRC Virology Unit, Glasgow, Scotland), and a third was that of the University of Wisconsin Genetics Computer Group (Deveraux et al. 1984).

Transcription of the cloned repeat in a cell-free system was studied according to the method of Manley et al. (1980). For discriminating between RNA polymerase II and III, transcription was performed either in absence or presence of α -amanitin (1 μ g/ml).

For the in situ hybridization analysis, long-term fibroblast cultures were used. Preparations were denatured in $0.1 \times \text{SSC}$ ($1 \times \text{SSC} = 0.15 \text{ M NaCl}$ and $0.015 \text{ M Na citrate}$) just below boiling temperature (Singh et al. 1977). A ^3H nick-translated (Rigby et al. 1977) recombinant plasmid was used as a labeled probe. C-banded preparations were from the same specimen.

Results

Sequencing of the killer-whale repeat was based on readings of both strands of the repeat. The sequencing strategy is shown in figure 1. The length of the repeat was 1,579 bp. The sequence of the repeat is given in figure 2. The central A of the *EcoRV* site was chosen as position 1 of the sequence. The G + C content of the repeat measured 44%. This figure is close to the G + C content of cetacean mainband DNA as judged from isopycnic ultracentrifugation (Árnason et al. 1977, 1978).

The 1,579-bp sequence was analyzed with respect to presence and distribution of repetitions having the same orientation. Five 9-bp sequences occurred twice in the repeat. In no case were these sequences adjacent to each other. The distribution of these sequences showed no particular pattern, and thus these identities did not indicate that the 1,579-bp repeat had evolved by means of an accumulation of short repeats in direct orientation.

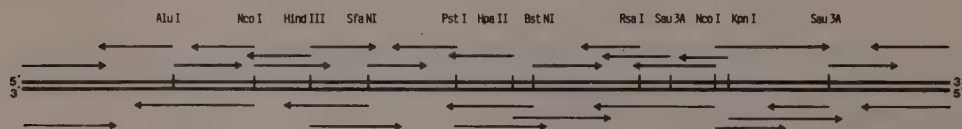


FIG. 1.—Sequencing strategy of the 1,579-bp repeat of the killer whale, *Orcinus orca*. Restriction sites used for labeling are indicated. Both strands of the repeat were sequenced according to the figure.


```

1  ATCTTATGATGTTCTTTTTTGTGTGACTTATTTTCATGTAGAATCATCGTACCTGAATC
61  CACTCATTATGCTGCTACGGGCTGATGACATAGATTTTCATGCTGAGTGATATGTCATT
121 GTACGTAAGTACCACAAAGCTCTTTATCCATTTTTCGCTTTTCGGCATCTGAAGTTGAC
181 CATAAAGCAGGTTCTGTAAACAGAGCCGTACAAACTTTGGGATGGCTGTGCTTTTTTG
241 ATTTTAATTTCAACTAAGCTATAGGACCATAAGTGGAAAGTGCCTAGGCTGTGGCTTT
301 GTTGTTTAGATGTTTCAGGAAACACCATACACTTCCAGAGTGGCTGTGGCAATTTAC
361 ATCCGGCCATCAGCATACAAGGCTCCCAATTCCTCCAGGCTGTCCCGCTTTCTGGA
421 TTTTACACTTTTTTCAGATGGCCCTTTGACCGTGGGGCAGTGAGACTTCATGATGCA
481 GATTTCTTTTCAAGCTTGCTTGGTGGCCAAAAAGTTCGTATGGCTTTCTTCTGAATA
541 TATTAGGAAAAACGCATAAGACCTTTTGGCCAAGTGCATCATTGGGACGTTCTGGCC
601 TGTTTTCTATGCTTTACATGCAATTCACGCTACCTCCTGAAATCGGTGTGCGCAATT
661 TTGCCCCGCTTCAAGTCTCTTGGCAGCCTTACTTCAATATATTTTGGACGATAGCTG
721 TCATTTATAACTCTGCAGGTTTGTGAATCACAGTGCCCTGAGCTCTTTCTTCAACTCG
781 CTTTCTTGTGAGCTGGCTGCAACACCGCAGGATTGCTTGAGGCCCTCGTGTGGTCCGGC
841 AGGCGACGCTGAGCCTTTTGTAACTCTCTTCCGTGGGAAATGAGGCTTAAATTTGG
901 CCGTCCAGACACCTCCAGCTAGTCTCTCATTGATCTCCCTATTCCTGTTTCTTTCCGC
961 AGAAATGCAAACTGGGCCAAACAGGAGGTTAAAGGCACTGACTCTCAAGTGGGGAGAG
1021 TGTTAGTAAAGCGCTGGAATGTTGCACCGAGTACCAGGGGACGAAACTGAGACACAT
1081 TTGAACACGTTTCCCGATCACACGGTGGATCATACTCGGTTCCACATGCATGTTTATAG
1141 CTGAAGGAAGAATCCCTTAAACCTGGAGAGTTGAGACCATGGAATGGGATACCATGCAAT
1201 ATGACTTCAAAAGGCTCTGATTGCTAACTGAAACTCACATACTATCACTGCTGGCTT
1261 TATGCCACTGTACACACGCTTGATTCCTTTTCGGAGACATATAAGTCCACAGGTTTTAAG
1321 ATTCCTACTAGTCAGGTATATTCTAGGAGTTAATATGGAGCTGTGAGCTCCACTCGCAT
1381 AGCAAGGAGTAGCTCTTGTCTATTAATATTGGCTTATGGAACGGTATCTGTGCTGATT
1441 TCAATCTCTGCTGTTTATGCAGCACCTCAACTCACCTTTCCCTTAAAGCAAGCAATGTTG
1501 GTGTTCTACATTGAGACCCCTGTTGTTTTGTAATTCAGTCTCTGTGATGCCAAGTTA
1561 CATTCCTGTAGTAGTAGT

```

FIG. 2.—Base sequence of the killer-whale 1,579-bp repeat.

The presence of inverted repeats within which intrastrand pairing might occur was also investigated. Several such regions of high quality (UWCGC stemloop program; Deveraux et al. 1984) were observed. The most conspicuous of these was a 72-bp-long sequence in position 505–576. The quality value for complementarity for this sequence was 76.0. The sequence was preceded in position 491–500 by a 10-bp sequence having full intrastrand complementarity. A 100-bp sequence (position 485–584) that included the two inverted repeats, displayed as cruciform structures, is shown in figure 3. Other inverted repeats having high-quality values were found in positions 201–233, 251–304, and 717–754 (fig. 4). The 251–304 region is interrupted in position 269–286 by a sequence without intrastrand complementarity.

Several regions with an open reading frame were found in the repeat. The longest segment was a 330-bp sequence in position 398–728. Other regions of substantial length were located in positions 522–747, 1,262–952, and 1,112–942 (fig. 4).

Transcription in a cell-free system (Manley et al. 1980) was recorded in the repeat. The length of the transcript was ~340 nucleotides (fig. 5). The results were repeatable in the sequenced repeat but could not be verified in other clones.

Screening for sequence similarities between the 1,579-bp repeat and primate Alu sequences revealed sequence similarities in some cases. The most conspicuous similarities were found in position 1,143–1,168 of the killer-whale repeat and in position 163–188 of a cloned Alu sequence (Rubin et al. 1980) in which a 26-bp sequence had 22 matches. An 18-bp sequence, position 807–824, showed 17-bp identities with position 298–315 of an Alu sequence of the African green monkey (Saffer and Lerman 1983). These similarities were not elaborated on further, since they might be coincidental.

A highly repetitive component of the bowhead (*Balaena mysticetus*), a whalebone whale, has been sequenced in most of its length. The length of the repeat was 1,746 bp. Comparison between the killer-whale 1,579-bp repeat and the bowhead repeat showed substantial similarities between the two components. Figure 6 shows graphically the results of an alignment of the two sequences. Positions 1,247–1,426 in the bowhead

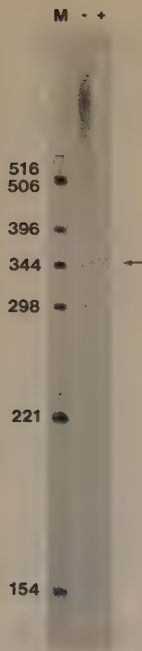


FIG. 5.—In vitro transcription of the sequenced killer-whale repeat. For discriminating between RNA polymerase II and RNA polymerase III activity, RNA synthesis was performed either in the absence (–) or in the presence (+) of α -amanitin. Electrophoretic separation was accomplished through a 6% polyacrylamide gel containing 8 M urea. A transcript, ~ 340 bp long, was recorded (arrow). M = pBR322 size marker.

component are missing in the killer-whale repeat. The distance between the *Pst*I and *Sac*I sites is the same in both repeats. Figure 7 shows an alignment of a 100-bp sequence, position 701–800, in which the *Pst*I and *Sac*I sites are located. Deviation from complete agreement was recorded in 14 positions.

C-bands are prominent in the killer-whale karyotype (Árnason et al. 1980). The C-bands are predominantly located in terminal and interstitial chromosome positions (fig. 8a). The chromosomal distribution of the 1,579-bp repeat was determined by in situ hybridization, the results of which are shown in figure 8b. The repeat was located in all major C-band-positive regions of the karyotype except in the long arm of chromosome sm9. Conspicuous C-band heteromorphism was observed in some chromosome pairs. The degree of heteromorphism was paralleled by different labeling intensity. This is particularly evident in the short arms of chromosomes st1 and st2 in figure 8b.

Discussion

Some cetacean species have been shown to harbor highly repetitive components with highly different degrees of divergence (Árnason and Widegren 1984). In the mysticete genus *Balaenoptera*, three unrelated, highly repetitive components have been studied. One variable component is specific for the genus. This component is assumed to be more recent than the other two, which occur outside the balenopterids. Of these two components, one characterizes the whalebone whales, and the other is the common 1,740-bp cetacean component. Conceivably, the presently analyzed 1,579-bp repeat

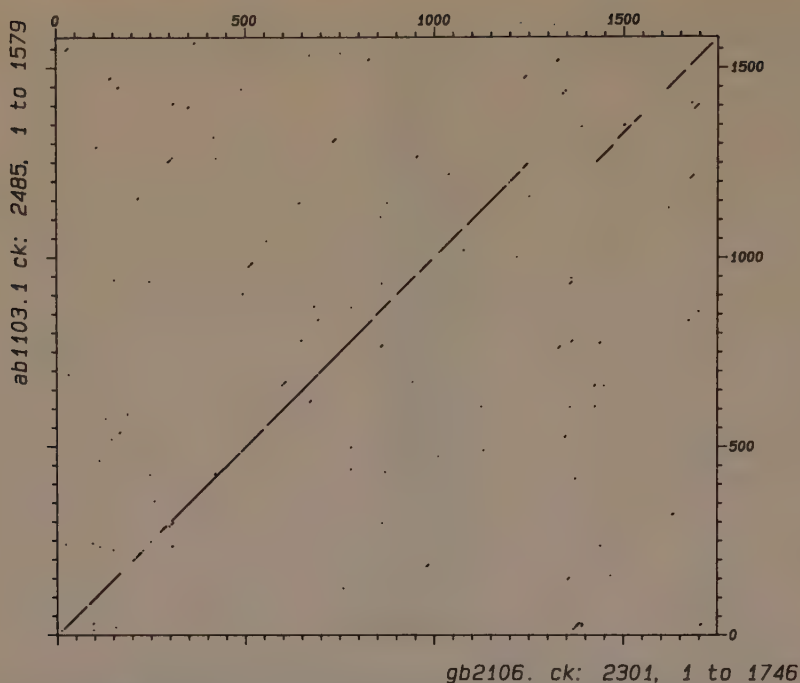


FIG. 6.—Graphic demonstration of results of alignment between the killer-whale, *Orcinus orca*, (vertical axis) and bowhead, *Balaena mysticetus*, (horizontal axis) repeats. Positions 1,533–1,605 of the bowhead have not been sequenced. The length of the killer-whale repeat is 1,579 bp, and the length of the bowhead repeat is ~1,740 bp. Region 1,247–1,426 of the bowhead component is missing in the killer-whale repeat. A dot is placed on the diagram whenever 14 of 21 consecutive aligned nucleotides match (Deveraux et al. 1984).

evolved from the 1,740-bp repeat that it has largely replaced in the delphinids. A common notion (e.g., Lewin 1982) is that highly repetitive DNA diverges rapidly, with each species having its own specific pattern of highly repetitive DNA. The 1,740- and 1,579-bp components differ radically from this concept, since their evolution is evidently very slow.

Molecular hybridization between various cetaceans has shown that the amount of the 1,740-bp repeat is limited in the family Delphinidae and that the delphinid repeat is absent in other cetacean families. It can be assumed that the delphinid repeat replaced the 1,740-bp repeat prior to the evolutionary radiation of the Delphinidae

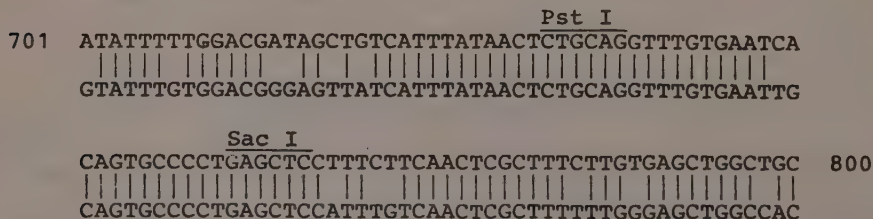


FIG. 7.—Alignment of 100-bp regions (position 701–800) of the killer-whale (top) and the bowhead (bottom) repeats. The distance between the *Pst*I and *Sac*I sites is the same in both repeats. Nonidentical nucleotides occurred in 14 positions.

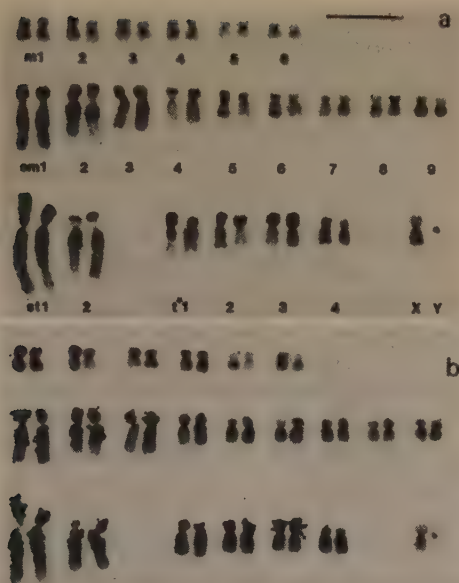


FIG. 8.—The karyotype of the killer whale. *a*, C-banded karyotype; *b*, results of in situ hybridization using a cloned repeat as labeled probe. The 1,579-bp component is present in all major C-bands. The C-band heteromorphisms in *a* are reflected by different degrees of labeling in *b*. This is particularly evident in the short arms of chromosomes st1 and st2. Bar = 10 μ m. Both cytological preparations are from the same specimen.

and that the repeat has been conserved during the later evolution of the delphinids, irrespective of the different degrees of amplification in different delphinid genomes. The greatest amplification has occurred in the killer-whale genome in which the number of repeats was estimated at $4\text{--}5 \times 10^5$, comprising more than 15% of the genome.

The results of the in situ hybridization showed that the amplification of the 1,579-bp repeat in the killer-whale genome had taken place both at ancestral delphinid C-band sites and also at C-band sites unique to the killer whale.

The way by which repetitive DNA elements spread through eukaryotic genomes is still not known in detail, but substantial information is available on transposable elements in *Drosophila* (Rubin 1983) and the Alu sequences in various mammals. Transposable DNA elements are often flanked by tandem repeats of DNA, but no such structures were found in the 1,579-bp repeat. These observations do not, however, exclude the possibility that such structures may exist in a limited number of repeats and may promote the introduction of the repeat at new chromosomal sites. It has been proposed (Jagadeeswaran et al. 1981; Van Arsdell et al. 1981) that transposition of Alu DNA may be directly linked to its transcription. Transcription was registered in the now sequenced 1,579-bp repeat, and a TGGCCAAGTG sequence corresponding to the consensus promoter sequence of RNA polymerase III (Galli et al. 1981) occurs in the 1,579-bp repeat. The 10-bp sequence is in position 570–579 of the repeat at the very end of the inverted 72-bp repeat. The sequence is within two open reading frames of substantial length. It is probable that the transcriptional activity observed in the repeat is connected with the presence of the RNA polymerase III promoter consensus sequence, but the localization and the general occurrence of transcriptional activity is yet to be determined.

Smith (1976) demonstrated that sister chromatid exchange would impose uniformity within clusters of highly repetitive DNA, either by eliminating new repeat varieties or by increasing their frequency in any particular chromosome region. The validity of Smith's (1976) model is easily conceived with respect to maintaining homogeneity within a repeat cluster, but it appears unlikely that it would prevent divergence between highly repetitive components residing on different chromosome pairs. Restriction analysis of the delphinid repeat in different species (Árnason et al. 1984) has shown that the component has been strikingly conserved during the 20–24 Myr of delphinid evolution, with respect to both the length of the repeat and the localization of various restriction sites. The conservation of the 1,740-bp repeat in other cetacean families is still more conspicuous, since this repeat has been maintained in lineages that separated >40 Myr ago.

The possible functions of highly repetitive DNA are still obscure. However, the presence in the same organism of highly repetitive DNAs with entirely different rates of evolution may indicate that different highly repetitive DNA components are subject to different selective pressures. In the balenopterid whales, one rapidly diverging and two conserved components have been studied in some detail (Árnason et al. 1978; Árnason and Widegren 1984). The rapidly diverging (light-satellite) component is present in centromeric C-bands in just a few chromosome pairs. Compared with the other two components, the light satellite appears to evolve quite freely and to be subject to limited selective pressure with respect to its composition. However, it does not follow from this that such components are immaterial to the organism that carries them. C-band heteromorphism is striking in most cetaceans (Árnason 1974), and based on these findings and the observations of Nankivell (1976) and Miklos and Nankivell (1976) that chiasmata positions are affected by the distance to C-band sites, Árnason et al. (1978) suggested that the different amounts of highly repetitive DNA in heteromorphic C-bands would put adjacent euchromatin out of parity and thus reduce the meiotic recombination in the vicinity of highly repetitive DNA. This effect would be achieved just by the presence of the heterochromatin, irrespective of the composition of the highly repetitive DNA.

The conservation of the presently studied delphinid repeat and the 1,740-bp cetacean repeat makes it highly plausible that these components are subject to selective pressure. Such pressure might be based on transcriptional activity in particular tissues or at certain developmental stages, or on structural properties of these components that might affect the spatial organization of the chromosomes in meiosis or mitosis. Evolutionary conservation of these components might also be imposed by their being selfish DNA (Orgel and Crick 1980) whose structure is optimal for being maintained and propagated in cetacean genomes. However, since it is hardly possible to test the selfish DNA hypothesis experimentally, we feel that possible functions of this repeat in cetacean genomes need to be thoroughly studied before other explanations are sought.

Acknowledgments

This work was supported by the Swedish Natural Science Research Council (grants B-BU 3655-105 and B-BU 3655-106) and by The Erik Philip-Sörensen Foundation.

LITERATURE CITED

- ÁRNASON, Ú. 1974. Comparative chromosome studies in Cetacea. *Hereditas* 77:1–36.
 ÁRNASON, Ú., M. HÖGLUND, and B. WIDEGREN. 1984. Conservation of highly repetitive DNA in cetaceans. *Chromosoma* 89:238–242.

- ÁRNASON, Ú., A. LIMA-DE-FARIA, H. GRANSTRÖM, and M. ISAKSSON. 1977. Analysis of cetacean DNA. Localization of 18S and 28S ribosomal RNA cistrons in a heavy DNA component. *Hereditas* **87**:67–76.
- ÁRNASON, Ú., R. LUTLEY, and B. SANDHOLT. 1980. Banding studies on six killer whales: an account of C-band polymorphism and G-band patterns. *Cytogenet. Cell Genet.* **28**:71–78.
- ÁRNASON, Ú., I. F. PURDOM, and K. W. JONES. 1978. Conservation and chromosomal localization of DNA satellites in balenopterid whales. *Chromosoma* **66**:141–159.
- ÁRNASON, Ú., and B. WIDEGREN. 1984. Different rates of divergence in highly repetitive DNA of cetaceans. *Hereditas* **101**:171–177.
- DEVERAUX, J., P. HAEBERLI, and O. SMITHIES. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**:387–395.
- GALLI, G., H. HOFSTETTER, and M. C. BIRNSTIEL. 1981. Two conserved sequence blocks within eukaryotic tRNA genes are major promotor elements. *Nature* **294**:626–631.
- JAGADEESWARAN, P., B. G. FORGET, and S. M. WEISSMAN. 1981. Short interspersed repetitive DNA elements in eucaryotes: transposable DNA elements generated by reverse transcription of RNA pol III transcripts? *Cell* **26**:141–142.
- LEWIN, R. 1982. Repeated DNA still in search of a function. *Science* **217**:621–623.
- MANLEY, J. L., A. FIRE, A. CANO, P. A. SHARP, and M. L. GESTER. 1980. DNA dependent transcription of adeno virus genes in a soluble whole cell extract. *Proc. Natl. Acad. Sci. USA* **77**:3855–3859.
- MAXAM, A. M., and W. GILBERT. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**:499–560.
- MIKLOS, G. L. G., and R. N. NANKIVELL. 1976. Telomeric satellite DNA functions in regulating recombination. *Chromosoma* **56**:143–167.
- NANKIVELL, R. N. 1976. Karyotype differences in the crenaticeps group of Atractomorpha (Orthoptera, Acridoidea, Pyrgomorphidea). *Chromosoma* **56**:127–142.
- ORGEL, L. E., and F. H. C. CRICK. 1980. Selfish DNA: the ultimate parasite. *Nature* **284**:604–607.
- RIGBY, P. W., M. DIECKMANN, C. RHODES, and P. BERG. 1977. Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. *J. Mol. Biol.* **113**:236–251.
- RUBIN, G. M. 1983. Dispersed repetitive DNAs in *Drosophila*. Pp 329–361 in J. A. SHAPIRO, ed. *Mobile genetic elements*. Academic Press, New York.
- RUBIN, G. M., C. M. HOUCK, P. L. DEININGER, T. FRIEDMANN, and C. W. SCHMID. 1980. Partial nucleotide sequence of the 300 nucleotide interspersed repeated human DNA sequences. *Nature* **284**:372–374.
- SAFFER, J. D., and M. I. LERMAN. 1983. Unusual class of Alu sequences containing a potential z-DNA segment. *Mol. Cell Biol.* **3**:960–964.
- SINGH, L., I. F. PURDOM, and K. W. JONES. 1977. Effect of different denaturing agents on the detectability of specific DNA sequences of various base compositions by in situ hybridization. *Chromosoma* **60**:377–389.
- SMITH, G. P. 1976. Evolution of repeated DNA sequences by unequal crossover. *Science* **191**:528–535.
- VAN ARSDELL, S. W., R. A. DENISON, L. B. BERNSTEIN, A. M. WEINER, T. MANSER, and R. F. GESTELAND. 1981. Direct repeats flank three small nuclear RNA pseudogenes in the human genome. *Cell* **26**:11–17.

KEN W. JONES, reviewing editor

Received February 14, 1985; revision received April 9, 1985.

Pinniped phylogeny enlightened by molecular hybridizations using highly-repetitive DNA

Ulfur Arnason and Bengt Widegren
Department of Genetics, University of Lund, Sweden

Key words: highly-repetitive DNA, pinniped evolution.

Address for correspondence and reprints: Dr. Ulfur Arnason, Department of Genetics,
University of Lund, Sölvegatan 29, S-223 62 Lund, Sweden.

Running head: highly-repetitive DNA and Pinniped Phylogeny

Abstract: Four non-cross-hybridizing, highly-repetitive DNA components of the Weddell seal were cloned and used for Southern blot hybridizations in order to clarify pinniped phylogeny. Each of the components was present and possessed the identical fragment length in all pinnipeds, representing true seals, walrus and sea lions. Three of the four components also hybridized to fragments of the same length in all mustelids except skunk. Limited hybridization also occurred to raccoon DNA. Hybridization to other terrestrial carnivores (polar bear, dog, cat) was barely recognizable. The fourth component was pinniped specific. The results are compatible with the mustelids and pinnipeds being sister groups. The findings also suggest that the pinnipeds are monophyletic, having separated from mustelid ancestors as one lineage that later differentiated into otariids, odobenids and phocids. The hybridizations indicated clear differences between the skunk and other mustelids.

INTRODUCTION

The application of molecular biological approaches to questions in molecular biology and phylogeny is becoming more widespread. The increasing interest stems from the fact that convergence in the evolution of DNA sequences is unlikely to occur. In population biology, analyses of mitochondrial DNA are rapidly replacing the use of protein polymorphisms, and mtDNA has been shown to be also suitable for obtaining information on species relationships within a particular genus (Brown 1983; Ferris et al. 1983). Molecular hybridizations based on single copy DNA have provided valuable information on phyletic relationships at various levels (e.g. O'Brien et al. 1985). A particular advantage of this approach is that long range relationships can be elucidated. However, as demonstrated by Templeton (1985), who scrutinized hominoid data, considerable care is needed when assessing phyletic affinities where successive divergence times are small compared to the total elapsed time.

Cladistic characters based on molecular hybridizations that were easily recognized over the range from genus to order would be valuable for studies of evolutionary biology and phylogeny. At least in some instances, highly-repetitive DNA appears to fulfill those requirements. So far, however, the use of it has been rather limited, probably because earlier reports proclaimed its species specificity, rather than its usefulness in study of common components in more distantly related taxa.

Studies of DNA in cetaceans have shown that some highly-repetitive components can maintain their characteristics through long periods of evolution, yielding unequivocal evidence of phyletic relationships, even at the subordinal level (Arnason 1982a; Arnason et al. 1984; Widegren et al. 1985).

In the present study we turned our attention to the other large group of marine mammals, namely the pinnipeds, with the following particular aims: (1) to investigate to what extent highly-repetitive DNA can be used to enlighten phyletic relationships that

have not been conclusively settled by other methods; (2) to answer the question whether the pinnipeds are mono- or diphyletic, and (3) to identify the closest relatives of pinnipeds among terrestrial carnivores. Clarification of these issues is of particular interest, because results based on immunological comparisons suggesting that the pinnipeds are monophyletic (Sarich 1969, 1975) are at variance with most paleontological, anatomical and zoogeographical studies (McLaren 1960; Mitchell and Tedford 1973; Tedford 1976; Ray 1976; Repenning and Tedford 1977; Janvier 1984). Karyological analyses on pinnipeds have emphasized the prominent karyological uniformity among the pinnipeds (Fay et al. 1967; Arnason 1974b, 1977, 1982b; Anbinder 1980) but have not answered conclusively the question of their ancestry.

MATERIALS AND METHODS

The pinniped DNA's used in the present study included Weddell seal (Leptonychotes weddelli), leopard seal (Hydrurga leptonyx), crabeater seal (Lobodon carcinophagus), bearded seal (Erignathus barbatus), hooded seal (Cystophora cristata), harbour seal (Phoca vitulina), large seal (Phoca largha), ringed seal (Phoca hispida), harp seal (Phoca groenlandica), walrus (Odobenus rosmarus), California sea lion (Zalophus californianus) and northern sea lion (Eumetopias jubatus). Of the true seals L.w., H.l., L.c., E.b., and C.c. have $2n=34$ chromosomes; P.v., P.l., P.h., and P.g. have $2n=32$. The terrestrial carnivore DNA's included the mustelids: otter (Lutra canadensis), ferret (Mustela putorius), mink (Mustela vison), European badger (Meles meles), striped skunk (Mephitis mephitis) and skunk (unidentified species). The non-mustelid terrestrial carnivore DNA's were raccoon (Procyon lotor), polar bear (Ursus maritimus) dog (Canis familiaris) and cat (Felis domesticus).

Preparation and purification of DNA was basically according to the method developed by Marmur (1961). The DNA of the California sea lion was prepared from placenta. This DNA had rather low molecular weight as the placenta was not as fresh as the other tissues used.

Four highly-repetitive DNA components were identified in the genome of the Weddell seal after electrophoretic separation of DNA restricted with Eco RI. The lengths of the components were approximately 2050 (A), 1650 (B), 900 (C) and 700 (D) base pairs (bp).

For cloning procedures, DNA was cleaved with Eco RI and separated on 1% preparative agarose gel. After staining with ethidium bromide, the components were dissected from the gel, electroeluted, ligated into the Eco RI site of the plasmid vector pUC8 (Vieira and Messing 1982) and cloned in *E. coli* K12-JM83. Positive clones were identified by colony hybridization (Grunstein and Wallis 1979).

DNA from the different species included in the comparison was restricted with Eco RI or Hae III. The digested DNA (1-2 µg) was loaded on 1 % agarose gels and separated overnight at low voltage (15-30 V). The gels were stained with ethidium bromide and photographed on a UV-transilluminator. The DNA size marker used was pBR322 cleaved with Rsa I.

Each of the four cloned highly-repetitive components of the Weddel seal was labeled with ^{32}P according to the method developed by Rigby et al. (1977) and used for Southern blot hybridizations. The four components used did not crosshybridize among themselves.

Transfers of DNA to nitrocellulose membranes and hybridizations were according to Southern (1975). Membranes were prehybridized at 64 °C in 3 x SSC (1xSSC is 0.15 M sodium chloride, 15 mM sodium citrate), 10 x Denhardt's (1966) solution with 100 µg/ml denatured herring sperm DNA. The same procedures and conditions were used for the hybridizations except that the concentration of Denhardt's solution was 2 x. Membranes were extensively washed in 0.5 x SSC, 0.2% sodium dodecyl sulfate at 64°C.

Autoradiography of the hybridized membranes was performed at -70 °C, using Kodak X-Omat RP film and intensifying screens. Time of exposure varied between 4 and 24 h. The autoradiograms were primarily scrutinized with respect to the presence or absence of fragments having the same length as the cloned components.

RESULTS

The results of hybridization using component A are shown in Fig. 1. Hybridization was primarily registered in a ≈ 2050 bp Eco RI fragment in all pinnipeds and all mustelids, except skunk. In the same materials hybridization also occurred in a larger fragment, probably a dimer of the 2050 bp fragment. In addition, hybridization in the pinnipeds was recorded in larger and smaller fragments, several of which had identical lengths in all species studied. Of the phocids included in Fig. 1, the Weddell, bearded and hooded seals have $2n=34$ chromosomes; the harbour seal has $2n=32$. Hybridization patterns of the Weddell seal and other Antarctic seals (leopard seal, crabeater seal) were indistinguishable and only the L.w. results were included. Similarly the harbour seal was used as a representative of the $2n=32$ phocids (ringed seal, large seal, harp seal) which all gave indistinguishable patterns of hybridization. The similarity between the hooded seal ($2n=34$) and the harbour seal ($2n=32$) was greater than between the harbour seal and any other $2n=34$ phocid.

Prolonged exposures allowed detection of decreasing degrees of hybridization to a ≈ 2050 bp fragment from the raccoon, striped skunk and polar bear genomes. No differences were observed between the two skunk materials and only the results of the striped skunk were included in the figures. Limited hybridization also occurred with cat DNA; in this species, however, the fragment length was different. Hybridization to distinct bands was not registered with dog DNA.

As mentioned in Materials and Methods the DNA of the California sea lion (*Z.c.*) had considerably lower molecular weight than other DNA samples used. In Fig. 1 and the other figures this presumably accounts for the hybridization smear in this species and the rather limited representation of larger fragments in comparison with other pinniped species included.

Component B was well represented in all pinniped genomes, digestion with Eco RI and Hae

III giving probe hybridizing fragments of identical lengths. Outside the pinnipeds, hybridization to discrete fragments was not registered after digestion with Eco RI. Digestion with Hae III gave identical fragment lengths in pinnipeds and mustelids (except striped skunk), Fig. 2a and b. The degree of hybridization was, however, considerably lower in the mustelids than in the pinnipeds under the conditions now applied. In previous experiments, a substantially higher degree of hybridization to mustelid DNA was obtained when less stringent conditions were used. This indicates differences in composition between component B in pinniped and mustelid genomes, although the Hae III fragment length has been preserved. Low degree of hybridization to other terrestrial carnivore DNA was recognized after prolonged exposures. A part of the component was maintained as a ≈ 1650 bp fragment in the raccoon, polar bear and dog. The relatively short Hae III fragment of the Weddell seal, to which component B hybridized was also a characteristic of other Antarctic phocids (leopard seal, crabeater seal) studied. This fragment did not occur in northern hemisphere phocids.

Component C hybridized with pinniped DNA but not with DNA of terrestrial carnivores, Fig 3. This component is thus a specific pinniped characteristic. Fragment lengths were identical in all species. Judged from the degree of hybridization, this component is less frequent in the pinniped genomes than the other components studied.

The results of a hybridization using component D as a labeled probe are shown in Fig. 4. Hybridization occurred both to the 700 bp fragment and to fragments of much larger size. The component was abundant in all pinnipeds. Outside the pinnipeds it was well represented in the mustelids (except skunk). Eco RI fragment lengths in pinnipeds and mustelids were identical.

DISCUSSION

As mentioned in the introduction, analysis of mtDNA and molecular hybridizations based on single copy DNA have great applicability, although at different levels, for answering phyletic questions. In situations where the resolution of mtDNA and single copy DNA analysis is not optimal, highly-repetitive DNA can be considered as a very valuable tool for phyletic comparisons.

The possible functions of highly-repetitive DNA are still obscure but the fact that a particular species may harbour several unrelated highly-repetitive components is an important feature that increases the power of resolution in phyletic studies, as some components may have common occurrence in the taxa compared, whereas others may be unique to a given lineage.

The abundance of highly-repetitive DNA makes it relatively easy to isolate from other parts of the genome using restriction endonucleases. This is of particular advantage when the highly-repetitive components can be used directly as labeled probes without first cloning them. In hybridizations used for phyletic studies, no significant differences would be expected between cloned and non-cloned probes, provided the isolated bands are composed of a single repetitive component. In Southern blot hybridizations, the appearance of discrete autoradiographic bands might facilitate the appreciation of results of molecular hybridizations for scientists not working in this field. The interpretation of the results from hybridizations using highly-repetitive DNA must, however, be made with care since the degree of hybridization will depend not only on the evolutionary distance but also on the copy number of the repeat. In the present study we therefore primarily considered the presence or absence of fragments with specific lengths.

In the present study, hybridization results appear to have provided conclusive answers to two questions in pinniped phylogeny; (a) which are the closest relatives of pinnipeds

among terrestrial carnivores?; and (b) have the pinnipeds evolved from these terrestrial relatives as one or two lineages?

These questions have been controversial, as different approaches have given different answers. As mentioned in the introduction, most paleontological, anatomical and zoogeographical information has been interpreted in favour of a diphyletic origin of pinnipeds. According to that view, phocids share common ancestry with mustelids, whereas odobenids and otariids have evolved from ursid ancestors. Immunological studies of albumins and transferrins (Sarich 1969, 1975) have shown a much closer relationship between phocids, odobenids and otariids than generally recognized by the advocates of a diphyletic pinniped origin. However, those results have made a limited impression on authors favoring a diphyletic pinniped evolution (e.g. Janvier 1984), probably because the most likely ancestral group among arctoid terrestrial carnivores could not be conclusively identified.

The earliest karyological comparisons of pinnipeds (Fay et al. 1967) could not point out the terrestrial ancestry of pinnipeds; nor could later studies, based on banded karyotypes, provide a definite answer, although great similarities were found between procyonid and otariid karyotypes (Arnason 1974b, 1977; Wurster-Hill and Gray 1975). The karyological studies clarified the relationships between the $2n=34$ and $2n=32$ chromosome phocid karyotypes, however, as well as the relationship between the $2n=36$ chromosome otariid and $2n=32$ chromosome odobenid karyotypes. Surveying the karyological studies of otariids and phocids, Anbinder (1980) advocated a monophyletic origin of pinnipeds, but he did not link their evolutionary ancestry directly with the mustelids.

In the work reported here, the four unrelated highly-repetitive DNA components studied were found in substantial amounts in the phocid, otariid and odobenid lineages, and their fragment lengths were strikingly well preserved. Two of the components, A and D, were well represented in the mustelid (except skunk) genomes, digestion with Eco RI yielding

identical basic fragment lengths in both groups. Component B did not cleave out as a discrete Eco RI fragment in mustelid genomes but Hae III gave identical fragment lengths in both pinnipeds and mustelids, although hybridization under the conditions used was considerably less in mustelids. Hybridization outside the mustelids was very limited, but the three components occurred with the same fragment length in the raccoon genome. Hybridization to skunk DNA was less than to raccoon DNA and hybridization to the other species was still more limited.

Component C had identical fragment length in all pinnipeds. This component was unobserved outside the pinnipeds. The specificity of component C may have been acquired in two ways, either by origination of the component in the pinniped lineage or by elimination of the component from the mustelid and other carnivore genomes. The first explanation is the more plausible in our opinion.

The present results indicate that all of the pinnipeds share a common and closer evolutionary ancestry with the mustelids than with any other group included in the comparison. An evolution of phocids from mustelid (lutrine) ancestry has been widely proposed. Evolution of otariids and odobenids from a mustelid rather than ursid-procyonid ancestry also is indicated by the present findings. The Enaliarctinae (Mitchell and Tedford 1973) have been regarded as the link between otarioids and an ursid ancestry. If the Enaliarctinae are the link between otarioids and their terrestrial arctoid ancestors, they must be a link between mustelids and pinnipeds rather than between ursids and pinnipeds. In our opinion the presence of a pinniped specific highly-repetitive DNA component suggests that the pinniped ancestors separated from the mustelids as a common lineage, and that the highly-repetitive component C proliferated in this lineage prior to the separation of the phocoid from the otarioid lineage.

As accounted for by Ray (1976), the evolution of otariids and phocids is separated with respect to both age and localization of early fossils. It is well substantiated that early

otariid evolution took place in areas surrounding the North-Pacific Ocean, whereas phocids evolved in the North Atlantic-Paratethyan region. The present results, which are compatible with a monophyletic origin of pinnipeds, do not necessarily contradict geographically separate origins of otariids and phocids provided they arose from a common lineage evolving from the mustelids and that this lineage was represented on both the Pacific and Atlantic sides of present Eurasia, alternatively on both sides of the present North-America.

It is likely that the evolution of both otariids and phocids was fast in the early stages of marine adaptation, hence evolutionary linkages with mustelids, based on paleontological findings, may be difficult to establish. The situation may be similar to that in Cetacea, in which the divergence into odontocetes and mysticetes has been difficult to substantiate by paleontological findings. Monophyletic origin was not established until chromosomal data became available (Arnason 1969, 1974). Results from later studies on cetacean highly-repetitive DNA confirmed the karyological findings (Arnason et al. 1984).

The prominent karyotypic uniformity that characterizes both pinnipeds and cetaceans has been related to limited inbreeding among these mammals in comparison with terrestrial mammals (Arnason 1972, 1974a,b, 1982b). The reproductive biology of these animals and their high mobility in an environment without distinct physical barriers are primary factors counteracting the establishment of reproductive isolates in which origination of new karyotypes would be promoted. Wilson et al. (1975) related slow karyotypic evolution to large body size, but this is not a particularly suitable explanation for marine mammals, since cetaceans may differ in size by a factor of 1000 or more without exhibiting any differences in rates of karyotypic evolution.

The present results show that conservation of the four highly-repetitive DNA components in pinnipeds coincide with slow karyotype evolution. These findings,

however, do not allow the conclusion that slow karyotypic evolution is necessarily linked with slow evolution of highly-repetitive DNA. In rorquals, genus Balaenoptera, both variable and highly conserved components are present, the variable one having originated the most recently (Arnason and Widegren 1984; Widegren et al. 1985). Those findings showed that there were different constraints on the evolution of different highly-repetitive components, their composition and repeat length being stringently conserved in some species but not in others. When the phyletic age of the pinnipeds is taken into account it becomes evident that the characteristics of the four components now studied have been well maintained.

Together with immunological findings (Sarich 1969, 1975), the present results indicate a single pinniped evolutionary lineage and suggest that the Pinnipedia should be maintained as a taxonomic unit, rather than placing the Phocidae in the Musteloidea, and Odobenidae and Otariidae in the Otarioidea as proposed by Tedford (1976) and Repenning and Tedford (1977).

It has been suggested (e.g. Ray 1976) that independent invasions from the north gave rise to the different Antarctic phocid species. The hybridization of component B to a Hae III fragment specific for the different Antarctic phocids instead suggests that differentiation into species took place in the Antarctic, as different invading populations would be unlikely to carry the same DNA constitution. That argument is not conclusive, however, if the Hae III fragment is a characteristic of the more ancient Monachus as well.

As mentioned in the introduction and earlier in the discussion highly-repetitive DNA has been shown to be useful for settling phyletic questions, primarily due to the presence of specific components in particular evolutionary lineages. The three particular objectives were given in the introduction. With respect to the three objectives the following conclusions can be made. (1) The present results confirmed the earlier findings, based on cetacean comparisons, that highly-repetitive DNA can provide answers to phyletic

questions even at the subordinal level. (2) The answer to the specific questions whether pinnipeds are mono- or diphyletic is that the general pattern of pinniped hybridization and the occurrence of a pinniped specific component suggest that the phocids, otariids and odobenids evolved as one lineage. (3) The molecular hybridizations suggested that the mustelids were the closest relatives to the pinnipeds among the terrestrial carnivores.

Besides yielding information on the pinnipeds, the molecular hybridizations showed striking differences between the skunk, genus Mephitis, and the rest of the mustelid group composed of the otter, ferret, mink and European badger. The highly-repetitive DNA components common to pinnipeds and the otter, ferret, mink and European badger, were barely recognizable in the skunk. Hybridization to the raccoon DNA was actually greater than to the skunk DNA. The results could be interpreted as the skunk being more distant from the other mustelids but the definite answer to this question must wait until reciprocal hybridizations have been performed.

FIGURE LEGENDS

- Fig. 1 Results of Eco RI digestion of DNA of pinnipeds and terrestrial carnivores (above). A, B, C and D show localization of highly-repetitive components of the Weddell seal used for Southern blot hybridizations. Abbreviations in all figures: L.w.: Weddell seal; E.b. bearded seal; C.c.: hooded seal; P.v.: harbour seal; O.r.: walrus; Z.c.: California sea lion; E.j.: northern sea lion; L.c.: otter; M.p.: ferret; M.v.: mink; M.m.: european badger; M.me: striped skunk; P.l.: raccoon; U.m.: polar bear; C.f.: dog; F.c.: cat. Marker DNA: pBR322 cleaved with Rsa I, size of fragments in kilo bases. Hybridization with component A (below) showed that it occurred with identical fragment lengths in all pinnipeds and all mustelids except the skunk. Hybridization of low degree was registered in the raccoon.
- Fig. 2 Component B was hybridized to DNA cleaved with Eco RI and Hae III respectively. Digestion with Eco RI (above) gave rise to discrete hybridization fragments in pinnipeds only. In pinnipeds, Hae III (below) produced the same fragment length as Eco RI. Hybridization to a discrete fragment also occurred in the mustelids (except skunk). Fragment lengths in pinnipeds and mustelids were identical.
- Fig. 3 Results of a hybridization using component C. This component is pinniped-specific and no hybridization occurred in any DNA of terrestrial carnivores. Fragment lengths were identical in all pinnipeds.

Fig. 4 Eco RI cleaved out component D with identical fragment lengths in all pinnipeds and mustelids, except skunk. Limited hybridization to a fragment of the same size was recorded in the raccoon. In the pinnipeds, hybridization to fragments of much larger size also occurred. The pattern was the same for all pinnipeds.

Acknowledgments

The authors express their sincere thanks to Dr's K. Benirschke and O. Ryder, San Diego Zoo, San Diego, USA; R. E. Benveniste, NCI, Frederick, Maryland, USA; W. N. Bonner and R. M. Laws, British Antarctic Survey, Cambridge, England; E. W. Born, Greenland Fishing Research, Copenhagen, Denmark; F. H. Fay, Institute of Marine Science, Fairbanks, Alaska, and B. O. Röken, Kolmården Zoo, Sweden, for the various samples of tissues and DNA. Prof. F. H. Fay's and Dr. C. A. Repenning's valuable comments on the manuscript are also acknowledged. The work was supported by the Swedish Natural Science Research Council (grants B-BU 3655-106 and B-BU 3655-107) and by The Erik Philip-Sörensen Foundation.

Literature cited

- Anbinder, E. M. 1980. The karyology and evolution of pinnipeds, Akademia Nauk, USSR (in Russian).
- Arnason, U. 1969. The karyotype of the fin whale. *Hereditas* 62:273-284.
- Arnason, U. 1972. The role of chromosomal rearrangement in mammalian speciation with special reference to cetacea and Pinnipedia. *Hereditas* 70:113-118.
- Arnason, U. 1974a. Comparative chromosome studies in cetacea. *Hereditas* 77:1-36.
- Arnason, U. 1974b. Comparative chromosome studies in Pinnipedia. *Hereditas* 76:179-226.
- Arnason, U. 1977. The relationship between the four principal pinniped karyotypes. *Hereditas* 87:227-242.
- Arnason, U. 1982a. Southern blot hybridization in cetaceans, using killer whale restriction fragment as a probe. *Hereditas* 97:47-49.
- Arnason, U. 1982b. Karyotype stability in marine mammals. *Cytogenet. Cell Genet.* 33:274-276.
- Arnason, U., M. Höglund, and B. Widegren. 1984. Conservation of highly-repetitive DNA in cetaceans. *Chromosoma* 89:238-242.
- Brown, W. M. 1983. Evolution of animal mitochondrial DNA. Pp 62-88 in Nei and Koehn eds. *Evolution of genes and proteins*. Sinauer, Sunderland, Mass.
- Denhardt, D. T. 1966. A membrane filter technique for the detection of complementary DNA. *Biophys. Res. Comm.* 23:641-646.
- Fay, F. H., V. R. Rausch, and E. T. Feltz. 1967. Cytogenetic comparisons of some pinnipeds (Mammalia: Eutheria). *Can. J. Zool.* 45:773-778.
- Ferris, S. D., R. D. Sage, E. M. Prager, U. Ritte, and A. C. Wilson. 1983. Mitochondrial DNA evolution in mice. *Genetics* 105:681-721.
- Grunstein, M. and J. Wallis. 1979. Colony hybridization. *Meth. Enzym.*, 68:379.
- Janvier, P. 1984. Cladistics: Theory, purpose, and evolutionary implications. Pp 39-75 in J. W. Pollard, ed. *Evolutionary Theory: Paths into the Future*. John Wiley and Sons Ltd.

Marmur, J., A procedure for isolation of deoxyribonucleic acid from microorganisms. J. Mol. Biol., 3, 208-218, 1961.

McLaren, I. A. 1960. Are the Pinnipedia biphyletic ? Syst. Zool. 9:18-28.

Mitchell, E. D., and R. H. Tedford. 1973. The Enaliarctinae, a new group of extinct aquatic Carnivora and a consideration of the origin of the Otariidae. Bull. Am. Mus. Nat.Hist.151:201-284.

O'Brien, S. J., W. G. Nash, D. E. Wildt, M. E. Bush, and R. E. Benveniste.1985. A molecular solution to the riddle of the giant panda's phylogeny. Nature 317:140-144.

Ray, E. R. 1976. Geography of phocid evolution. Syst. Zool. 25:391-406.

Repenning, C. A., and R. H. Tedford. 1977. Otarioid seals of the Neogene. US Geol. Survey Professional paper 992.

Rigby, P. W., M. Dickermann, C. Rhodes, and P. Berg. 1977. Labeling deoxyribonucleic acid into high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol. 113:236-251.

Sarich, V. M. 1969. Pinniped phylogeny. Syst. Zool. 18:416-422.

Sarich, V. M. 1975. Pinniped systematics: Immunological comparisons of their albumins and transferrins. Amer. Zool. 15:826.

Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98:503-507.

Tedford, R. H. 1976. Relationship of pinnipeds to other carnivores (Mammalia). Syst. Zool. 25:363-374.

Templeton, A. R. 1985. The phylogeny of hominoid primates: a statistical analysis of the DNA-DNA hybridization data. Mol. Biol. Evol. 2:420-433.

Vieira, J., and J. Messing. 1982. The pUC plasmids, a M13mp7 derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene 19:259-268.

Widegren, B., U. Arnason, and G. Akusjärvi. 1985. Characteristics of a conserved 1579-bp highly-repetitive component in the killer whale, Orcinus orca. Mol. Biol. Evol. 2:411-419.

Wilson, A. C., G. L. Bush, S. M. Case, and M-C King. 1975. Social structuring of mammalian populations and rates of chromosomal evolution. Proc Natl Acad Sci 72:5061-5065.

Wurster-Hill, D. H., and C. W. Gray. 1975. The interrelationships of chromosome banding patterns in procyonids, viverrids and felids. Cytogenet. Cell Genet. 15:306-331.

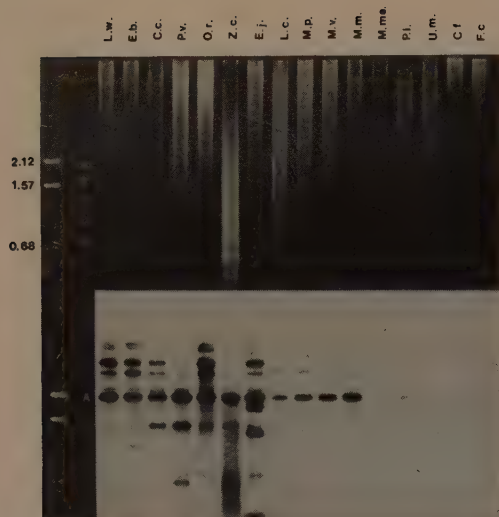


Fig. 1 Results of Eco RI digestion of DNA of pinnipeds and terrestrial carnivores (above). A, B, C and D show localization of highly repetitive components of the Weddell seal used for Southern blot hybridizations. Abbreviations in all figures: L.w.: Weddell seal; E.b. bearded seal; C.c.: hooded seal; P.v.: harbour seal; O.r.: walrus; Z.c.: California sea lion; E.j.: northern sea lion; L.c.: otter; M.p.: ferret; M.v.: mink; M.m.: european badger; M.me: striped skunk; P.l.: raccoon; U.m.: polar bear; C.f.: dog; F.c.: cat. Marker DNA: pBR322 cleaved with Rsa I, size of fragments in kilo bases. Hybridization with component A (below) showed that it occurred with identical fragment lengths in all pinnipeds and all mustelids except the skunk. Hybridization of low degree was registered in the raccoon.

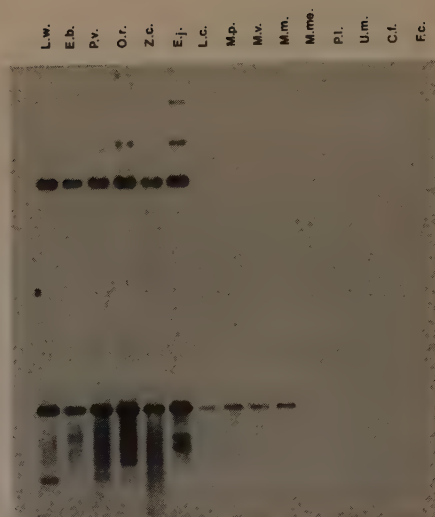


Fig. 2 Component B was hybridized to DNA cleaved with Eco RI and Hae III respectively. Digestion with Eco RI (above) gave rise to discrete hybridization fragments in pinnipeds only. In pinnipeds, Hae III (below) produced the same fragment length as Eco RI. Hybridization to a discrete fragment also occurred in the mustelids (except skunk). Fragment lengths in pinnipeds and mustelids were identical.

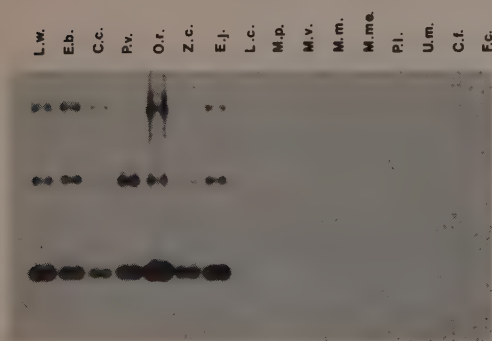


Fig. 3 Results of a hybridization using component C. This component is pinniped-specific and no hybridization occurred in any DNA of terrestrial carnivores. Fragment lengths were identical in all pinnipeds.

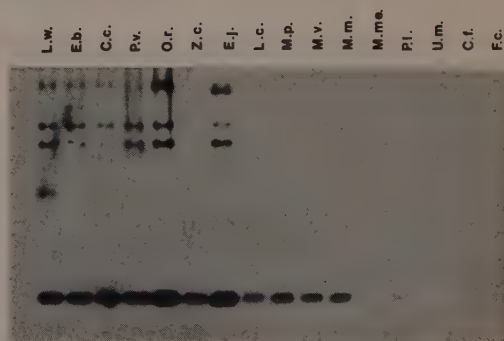


Fig. 4 Eco RI cleaved out component D with identical fragment lengths in all pinnipeds and mustelids, except skunk. Limited hybridization to a fragment of the same size was recorded in the raccoon. In the pinnipeds, hybridization to fragments of much larger size also occurred. The pattern was the same for all pinnipeds.

Analysis of the 420 bp highly repetitive component in balenopterid whales

Bengt Widegren and Ulfur Arnason

Department of Genetics, University of Lund, Sweden

Key words: highly repetitive DNA, whales, DNA sequencing, evolution

Running head: Sequence uniformity of highly repetitive DNA

Abstract

A highly repetitive component was isolated as a Xba I restriction fragment from DNA of three species of balenopterid whales. The component was cloned and sequenced. The length of the component was approximately 420 bp. The organization of the component was the same in the three species. About 220 bp comprised an unique sequence and the remaining part was made up of a number of repeats, six bp in length. Seven different clones of the component of the sei whale were sequenced. The sequences of one component from each of the sei and the fin whales were also determined. A consensus sequence for the nine components was calculated. In the unique part of the component, the sei whale sequences differed from the consensus sequence by an average of 2%. The corresponding figures for the minke and fin whales were 4.5 % and 2% respectively. Limited hybridization with DNA of some toothed whales was registered, however, the organization of the component was different in these species. The balenopterid characteristics of the component have been maintained during evolution of the genus Balaenoptera. This genus differentiated some 7-10 million years ago; and when this period of time is taken into consideration, it is apparent that the evolution of the component is very slow. The telomeric localization of the component in the balenopterid karyotypes and the conservation of the component indicate that the component may have a function in chromosome nuclear membrane interactions. Low degree of hybridization to the DNA of the chicken was also recorded.

INTRODUCTION

A considerable accumulation of knowledge on highly repetitive DNA has developed during the last decade. However, the biological functions of this type of DNA have not yet been resolved. There have been several reports (e.g. Jones and Singh 1982, Roizes and Pages 1982, Epple et al. 1983, Anachkova et al. 1984, Sanzo et al. 1984, Siddiqui et al 1984, Heller et al. 1984 and Anchkova et al. 1985) proposing possible functions of highly repetitive DNA, but so far none have conclusively pointed out functions which might explain the abundance of this type of DNA in the eucaryotic genomes. It is quite plausible that highly repetitive DNA originated and proliferated according to the principle of selfish DNA (Orgel and Crick, 1981) but the conservation of some highly repetitive sequences may indicate that particular functions evolved later. In order to obtain a deeper understanding of the possible functions that different classes of highly repetitive DNA may have and of the mechanisms by which they spread in the genome, it is essential that more data become available in this field.

The order Cetacea consists of two major lineages, namely, Odontoceti (toothed whales) with six extant families, and Mysticeti (whalebone whales) with three. The separation of the two lineages took place about 40 million years ago. Marine mammals in general and cetaceans in particular are characterized by strikingly slow karyotype evolution, presumably reflecting the relative absence of reproductive isolates among these animals, due to their good mobility in a homogenous environment without distinct physical barriers, and their slow rate of reproduction (Arnason 1972). These factors make cetaceans particularly attractive for studies of highly repetitive DNA as deviations from evolutionary patterns caused by the effects of reproductive isolates are unlikely to occur.

The present study includes an analysis of a highly repetitive component originally isolated as a heavy DNA satellite from balaenopterid whales (Arnason et al 1978) This

component was later cloned as a Xba1, 420 bp fragment from three species of the genus Balaenoptera. The evolution of the component is strikingly conservative and differences within and between species of the same genus are very limited, although the evolutionary separation of the three species studied occurred some 7-10 million years ago.

MATERIALS AND METHODS

The DNAs used were prepared according to the method of Marmur (1961). Xba1 fragments isolated from preparative gels were made blunt end using Klenow enzyme and cloned in the Eco RV site of pBR322. Cloning was also performed in the Xba1 site of M13 mp10. Both strands were sequenced following the Maxam and Gilbert (1980) procedures after 3' and 5' end labeling. Sequencing was also carried out following the procedure of Sanger (1981) after cloning in the Xba1 site of the vector M13 mp10.

DNA transfers and hybridizations were in accordance with Southern (1975). Probes were labeled with ^{32}P by nick translation (Rigby et al. 1977). Hybridizations were carried out in 3 x SSC at 64 °C , 5 x Denhardt's solution (Denhardt, 1966) with denatured bacterial DNA at a concentration of 100 µg/ml. Washings were at the same temperature in 1 x SSC and 0.5% SDS with several changes.

RNA was isolated from the testis, liver, spleen, brain and pancreas of the minke whale according to Pritchard and Holland (1985). After centrifugation through a CsCl cushion, the RNA was treated with DNase (Wellcome DPFF) in the presence of RNAsine. RNA (5 - 10 µg) was loaded on formaldehyde agarose gels (Maniatis 1982) and transferred to nylon membranes (GeneScreen). Hybridizations were performed with a ^{32}P labeled cloned component of the sei whale in the presence of bacterial DNA under the same conditions as for DNA hybridization.

Computer analyses of sequenced repeats were carried out using the University of Wisconsin Genetics Computer Group program (Deveraux et al. 1984) and the EMBL database, version number 5 (April 1985).

RESULTS

The composition of the Xba I component in three species of the genus Balaenoptera was determined by sequencing. The component which is tandemly organized (Arnason and Widegren 1984b) is about 420 bp long in the three species. The result of the sequencing showed that the component was composed of one unique part and another part made up of a number of repeats 6 bp in length. The length of the unique part was close to 220 bp in the three species. The composition of the 6 bp repeat was 5' GGGTTA 3' (5' TAACCC 3'). The general organization of the component is shown in the dot plot in Fig. 1.

The compositions of the sequenced components are given in Figs. 2. Deviation from the consensus sequence is indicated by lower case letters. The agreement between the unique parts of the components is striking but the number of the 6 bp repeats shows some variation. In Fig. 2 this variation has been accumulated at the end of the sequences. In the sei whale, the differences from the consensus sequence in the unique part of the components averaged 2.2%. The corresponding figures for the minke and fin whales were 4.7% and 2% respectively. In the sei whale, the GC content of the non repetitive was 40.3% and for the repetitive part it was 55% .In the other two species, the figures was close to those of the sei whale.

The DNA hybridization studies showed that the components occurred in other cetacean families, although the organization of the repeat was different from that of the balenopterids. The component was found to be particularly abundant in the right whale. Both the repeated (Fig. 3a) and the unique part (Fig. 3b) of the balenopterid component hybridized to the right whale DNA, giving the same pattern of hybridization in both instances. This indicates that the general organization of the component in the right whale is similar to that of the balenopterids although the basic unit is longer. Among the odontocetes the pattern was quite different. Neither part of the component was detectable in the La Plata dolphin (River dolphins). In the beluga whale (Narwhals) both the repetitive and the unique part of the component were present, but the organization of

both differed from that of the balenopterids. In the killer whale (Dolphins), only the unique part of the component was present at a detectable level. Hybridization with the unique part of the component revealed the presence of a tandem repeat with a basic unit length of approximately 200 bp. In the goosebeak whale (Beaked whales), hybridization to either part of the component was barely detectable. The repeated part of the component hybridized to a limited degree to sperm whale DNA but in this species the organization differed from that of mysticetes and other odontocetes. The unique part of the Xba I component also hybridized to sperm whale DNA. The limited hybridization was registered as a faint smear with fragment sizes shorter than 300 bp. Besides the cetaceans, limited hybridization was detected in some other mammalian species, notably the walrus, red deer and elephant (not shown). Besides the mammals, some hybridization homologies with chicken DNA were detected. The repetitive part showed a hybridization smear (Fig 3a) whereas the unique part revealed discrete bands at low intensity (Fig. 3b).

Open reading frames were found in the Xba I component. Northern blot hybridization (Alwine, 1980) to RNA from liver, spleen, pancreas, testis and brain, using cloned Xba I component as a labeled probe, however, did not indicate transcriptional activity in the tissues studied.

Computer comparisons between the Xba I component and the EMBL nucleotide data bank revealed similarities between the repetitive part of the component and two avian sequences. Both were globin genes from chicken and duck respectively. The sequence similarities between the whale repetitive component and the avian genes were located immediately upstream the first globin exon of the chicken (Dodgson and Engel 1983) and the duck (Erbil and Niessing 1982). Sequence homologies were also recorded with the Trypanosoma brucei tandem telomeric repeat (Blackburn and Challoner 1984). Dot plots of these homologies are shown in Fig.4.

Discussion

The sequencing data of the Xba I component in the three balenopterid species showed that the component had been strikingly well conserved during the evolution of genus Balaenoptera. The degree of hybridization to right whale DNA indicated that in this species, the amount of the component was similar to that in the balenopterids, although the organization of the component was different in the two mysticete families. Comparisons with some odontocete species showed that a relative to the balenopterid component was present in some species as well. However, the organisation of the component was different and the copy number much lower. The proliferation of the component appears to have coincided with the evolution of mysticetes and the manifestation of the characteristics of the 420 bp Xba I component presumably occurred together with the evolution of the genus Balaenoptera.

In the balenopterids, another component unrelated to the Xba I component, has been isolated and used for molecular hybridizations (Arnason et al. 1978; Arnason and Widegren 1984b). This component, originally isolated as a light DNA satellite, does not hybridize to DNA from species other than the genus Balaenoptera and it is therefore conceivable that it is younger than the the Xba I component. The light satellite shows considerable differences between the same three species included in the present study. In relation to the light satellite it is thus evident that the evolution of the Xba I component is very slow. The evolution of the Xba I component also appears to be slower than the evolution of nontranslated ribosomal DNA in cetaceans (Arnason and Widegren 1984a).

The comparisons between the three species showed that both the unique and the repetitive parts of the component were strikingly conserved. In the repeated part of the component there were some differences in the number of the six bp repeat unit but the composition of these units was basically the same. Individual sequences of the unique part of the component showed striking similarities both within and between species. Thus within species, the seven sequenced components of the sei whale differed, on

average, by only 2% from the consensus sequence. The same figure was obtained for the fin whale component. For the minke whale the figure was 4.5%. These figures give a divergency rate well below 0.5% per million years. Although not directly comparable, the corresponding figure for mtDNA is usually given as 2%. It is evident from the slow rate of evolution of the Xba I component that it does not conform with the previous notion that highly repetitive DNA diverges rapidly and therefore is species specific.

Provided the similarities between the Xba I component and avian DNA are orthologous and not paralogous it is evident that at least a part of the component evolved at a very early date. No hybridization to Trypanosoma DNA has been performed and the sequence homology with this species has not been investigated further.

In our opinion it is quite plausible that the amplification of the Xba I component followed the model of selfish DNA as proposed by Orgel and Crick (1981) but the conservation of this component in the balenopterid lineage indicates that the component in these animals has acquired particular biological functions. In this context the terminal localization of the component in the balenopterid and right whale karyotypes (Arnason et al. 1978, 1982) is worth considering as it is possible that the component is important for interactions with the nuclear membrane.

In our opinion it is quite plausible that the amplification of the Xba I component followed the model of selfish DNA as proposed by Orgel and Crick (1981) but the conservation of this component in the balenopterid lineage indicates that the component in these animals has acquired particular biological functions. In this context the terminal localization of the component in the balenopterid and right whale karyotypes (Arnason et al. 1978, 1982) is worth considering as it is possible that the component is important for interactions with the nuclear membrane.

Literature cited

Anachkova, B., Todorova, M., Vassilev, L. and Russev, G., Isolation of short interspersed repetitive DNA sequences present in the regions of initiation of mammalian DNA replication. *Eur. J. Biochem.* 141, 105-108, 1984.

Anachkova, B., Russev, G. and Altmann, H., Identification of the short dispersed repetitive DNA sequences isolated from the zones of initiation of DNA synthesis in human cells as Alu-elements. *Biochem. Biophys. Res. Com.* 128 (1), 101-106, 1985.

Alwine, J.C., Kemp, D.J., Parker, B.A., Reiser, J., Renart, J., Stark, G.R. and Wahl, G.M., Detection of specific RNAs or specific fragments of DNA by fractionation in gels and transfer to diazobenzoyloxymethyl paper. *Meth. Enzym.*, 68, 220, 1980.

Arnason, U., The role of chromosomal rearrangement in mammalian speciation with special reference to Cetacea and Pinnipedia. *Hereditas*, 70, 113-118, 1972.

Arnason, U., Purdom, I.F. and Jones, K.W., Conservation and chromosomal localization of DNA satellites in balenopterid whales. *Chromosoma*, 66, 141-159, 1978.

Arnason, U. and Widegren, B., Studies on ribosomal DNA in cetaceans (whales). *Hereditas*, 101, 149-154, 1984a.

Arnason, U. and Widegren, B., Different rates of divergence in highly repetitive DNA of cetaceans. *Hereditas*, 101, 171-177, 1984b.

Blackburn, E.H. and Challoner, P.B., Identification of a telomeric DNA sequence in *Trypanosoma brucei*. *Cell*, 36, 447-457, 1984.

Denhardt, D.T., A membrane filter technique for the detection of complementary DNA. *Biophys. Res. Commun.*, 23, 641-646, 1966.

Deveraux, J., Haeberli, P. and Smithies, O., A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acid Res.* 12, 387-395, 1984.

Dodgson, J.B. and Engel, J.D., The nucleotide sequence of adult chicken alpha-globin genes. *J. Biol. Chem.* 258, 4623-4629, 1983.

Epple, J.T., Cellini, A., Shorte, M. and Ohno, S., On evolutionary conserved simple repetitive DNA sequences: Do "sex-specific" satellite components serve any sequence dependent function? *Differentiation*, 23(suppl.) 60-63, 1983.

Erbil, C. and Nissing, J., The complete nucleotide sequence of the duck alpha-globin gene. *Gene*, 20, 211-217, 1982.

Heller, D., Jackson, M. and Leinwand, L., Organization and expression of non-Alu family interspersed repetitive DNA sequences in the mouse genome. *J. Mol. Biol.* 173, 419-436, 1984.

Jones, K.W. and Singh, L., Conserved sex-associated repeated DNA sequences in vertebrates. *Genome evolution (monogr.)* ed. Dover and Flavell, 135-154, 1982.

Maniatis, T., Fritsch, E.F. and Sambrook, J., *Molecular Cloning*. Cold Spring Harbor Laboratory, 1982.

Marmur, J., A procedure for isolation of deoxyribonucleic acid from microorganisms. *J. Mol. Biol.*, 3, 208-218, 1961.

Maxam, A.M. and Gilbert, W., Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.*, 65, 499-560, 1980.

Orgel, L.E. and Crick, F.H.C., Selfish DNA: the ultimate parasite. *Nature*, 284, 604-607, 1980.

Pritchard, R.H. and Holland, I.B. (ed), *Basic Cloning Techniques*. (Monogr.) Blackwell Scientific Publications, 6-9, 1985.

Rigby, P.W., Dieckmann, M., Rhodes, C. and Berg, P., Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. *J. Mol. Biol.*, 113, 236-251, 1977.

Roizes, G.P. and Pages, M., Conserved Sequences of bovine satellite DNAs. *Genome evolution (monogr.)* ed. Dover and Flavell, 94-111, 1982.

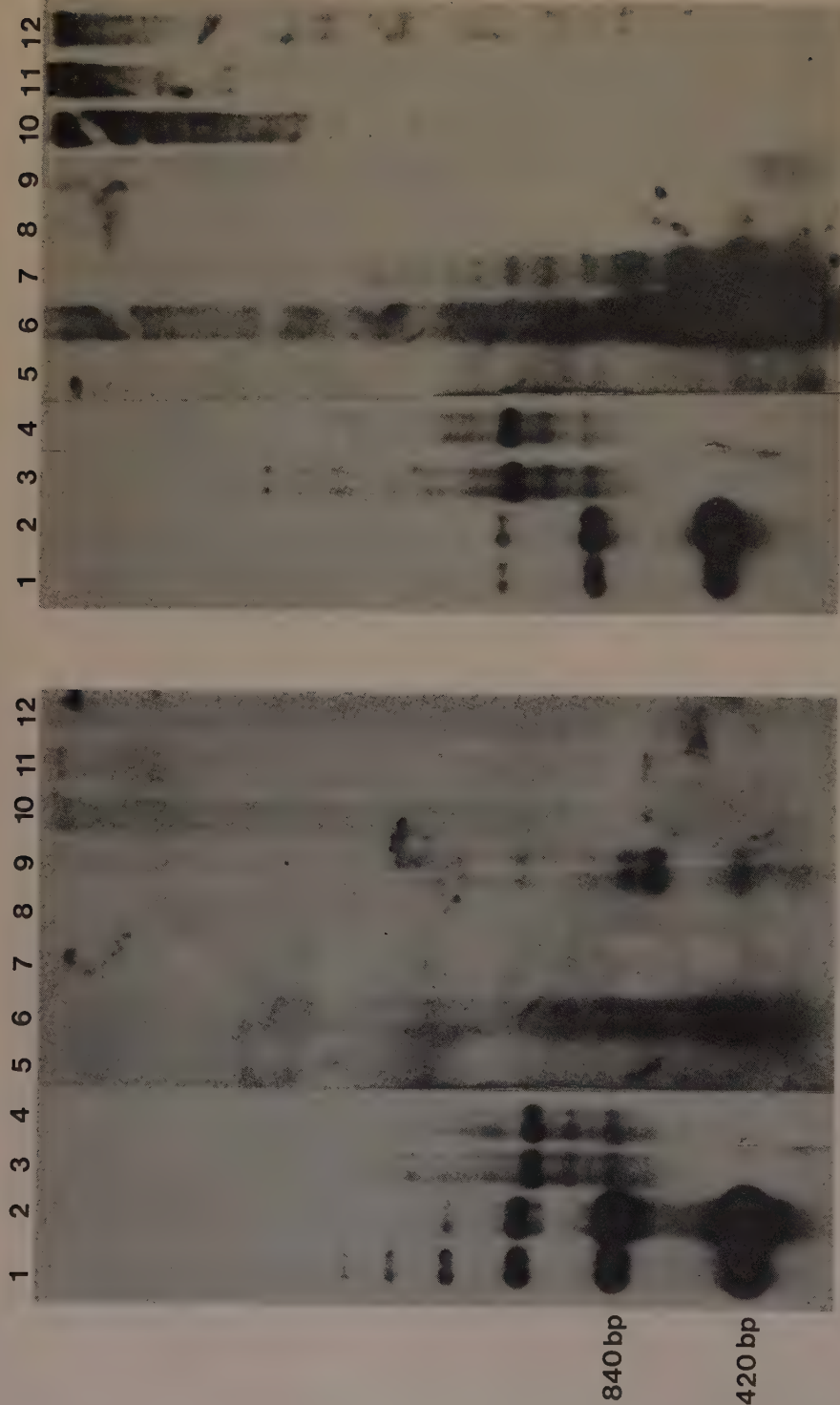
Sanger, F., Determination of nucleotide sequences in DNA. *Science*, 214, 1205-1210, 1981

Sanzo, M., Stevens, B., Tsai, M-J. and O'Malley, W.O., Isolation of a protein fraction that binds preferentially to chicken middle repetitive DNA. *Biochemistry* 23, 6491-6498, 1984.

Siddiqui, M.A.Q., Khandekar, P., Krauskopf, M., Mendola, C., Zarraga, A.M. and Saidapet, C., Control of muscle gene expression: Cardiac 7S RNA is homologous to 3'-noncoding Myosin gene and repetitive DNA sequences. *Expl Biol. Med.*, 9, 269-276, 1984.

Southern, E.M., Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol., 98, 503-517, 1975.

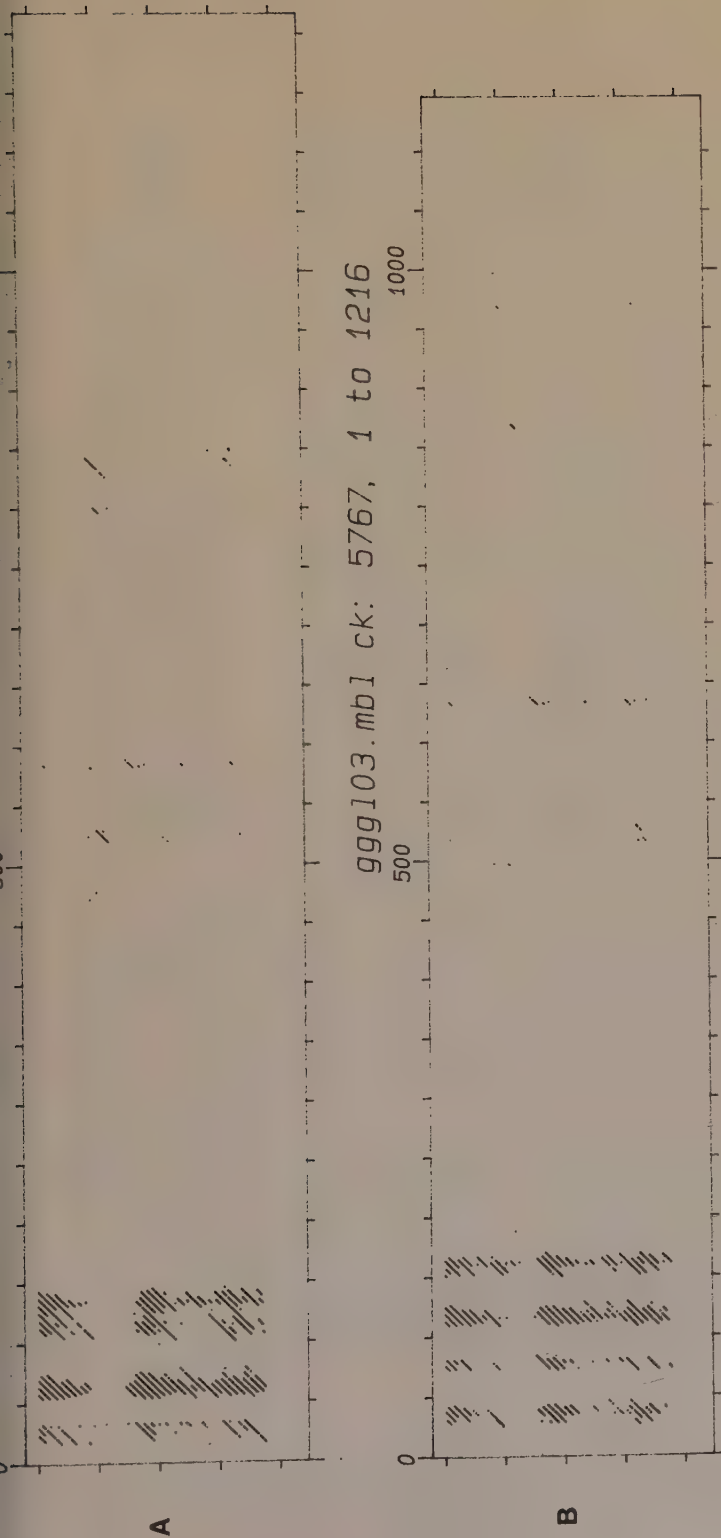
Widegren, B., Arnason, U. and Akusjärvi, G., Characteristics of a conserved 1579 bp highly repetitive component in the killer whale, *Orcinus orca*. Mol. Biol. Evol. 2(5), 411-419, 1985.



—A—

Figure 3. Results of the hybridizations with the repetitive part (A) and the unique part (B) of the 420 bp *Xba* I highly repetitive component from the minke whale. Number 1 to 10 in both (A) and (B) represent: 1: fin whale (B.p.) cut with *Xba* I, 2: minke whale (B.a.) cut with *Xba* I, 3 and 4: right whale (B.m.) cut with *Xba* I (3) and *Sty* I (4), 5: La Plata dolphin (P.b.) cut with *Rsa* I, 6: white whale (D.I.) cut with *Rsa* I, 7: killer whale (O.o.) cut with *Rsa* I, 8: goosebeak whale (Z.c) cut with *Rsa* I, 9: sperm whale (P.m) cut with *Rsa* I, 10 to 12: chicken (G.g.) cut with *Rsa* I (10), with *Bst* NI (11) and with *Alu* I (12) respectively. The figures 4A and B are composed from different exposures of each filter.

—B—



adhbada2.mbl ck: 1348, 1 to 1145

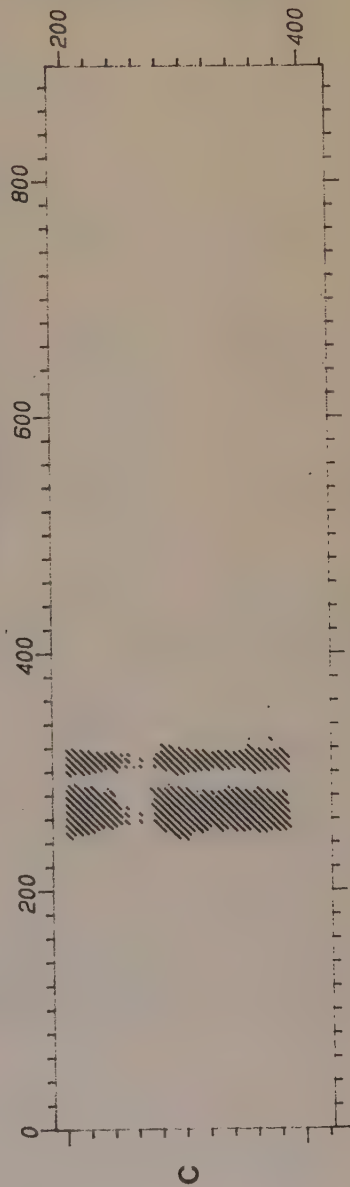


Figure 4. Dotplots, which graphically demonstrate the homologies between the repetitive part of the 420 bp Xba I component (reverse positions 420-190) and (A) the chicken α -globin gene with the 5' flanking region included, (B) the duck α -globin gene with the 5' flanking region included and (C) the trypanosoma telomeric highly tandemly repeated DNA.

trvsgrs.mbl ck: 7834, 1 to 897

B.b.1	1	TCTAGATACC	CCTTAGCAGC	TAAAGGAATA	TTATTTGGGG	GTCCAGCCCAT	GGAGGAGAGT	TTAaACACTA	GGATAGAGTA	AGGAACACAC	CCATTATAAA	100
B.b.2		TCTAGATACC	aCTTAGCAGC	TAAAGGAATg	TTATTTGGG	GTCCAGCCCAT	GGAGAGAGT	TTAGACACTA	GGATAGAGa	AGGAACACAC	CCgTTATAAA	
B.b.3		TCTAGATACC	CCTTAGCAGC	TAAAGGAATA	TTATTTGGGG	GTCCAGCCCAT	GGAGAGAA	TTAGACACTA	GGATAGATA	AGGACATAC	CCATTATAAA	
B.b.4		TCTAGATACC	CCTTAGCAGC	TAAAGGAATA	TTATTTGGGG	GTCCAGCCCAT	GGAGAGAGT	TTAGACACTg	GGATAGATA	AGGAACACAC	CCATTATAAA	
B.b.5		TCTAGATACC	CCTTAGCAGC	TAAAGGAATA	TTATTTGGGG	GTCCAGCCCAT	GGAGAGAGT	TTAGACACTA	GGATAA.ATA	AGGAACACAC	CCATTATgAA	
B.b.6		TCTAA. TACC	CCTTAGCAGC	TAAAGGAATA	TTATTTGGGg	GTCCAGCCCAT	GGAGAAAGAGT	TTAGACACTA	GGATAGATA	AGGAACACAC	CCATTATAAA	
B.b.7		TCTAGATACC	CCTTAGCAGC	TAAAGGAATA	TTATTTGGGG	GtgcAGCCCAT	GGAGAAAGAGT	TTAGACACTA	GGATAGATA	AGGAACACAC	CCATTATAAA	
B.a.		TCTAGATACC	CCTTAGCAGC	TAAAGGAATA	TTATTTGGGG	GtTcAGCCCAT	GGAGAAAGAGT	TTAGACACgA	GGAGAGATA	AGGAACACAC	CCATTATCAA	
B.p.		TCTAGATACC	CCTTAGCAGC	TAAAGGAATA	TTATTTGGGG	GtTCCAGCCAT	GGAGAGAGT	TTAGACACTA	GGATAGATA	AGGAACACAC	CCATTATAAA	
B.b.1	101	GAAATCACAT	TAGGATTCTC	TTTTTAAGCT	GTTTCCTTAA	ACACTAGTGC	TTAGGAAaC	TATTGAGGC	AGAaCAGTC	AAGGTAGCC	TAGGTTAGG	200
B.b.2		GAAATCACAT	TAGGATTCTC	TTTTTAAGCT	GTTTcTgAAA	ACACTAGTGC	TAAAGAAATC	TATTAGAGC	AGAAGAGTC	AAGGTAGCT	TatGGTTAGG	
B.b.3		GAAATCACAT	TAGGATTCTC	TTTTTAAGCT	GTTTgcTTAA	ACACTAGTNN	NNNNNNNN	NNNNNNNN	NNNNNNATC	AAGGTAGCC	TAGGTTAT.	
B.b.4		GAAATCACAT	TAGGATTCTC	TTTTTAAGCT	GTTTCCTTAA	ACACTAGTGC	TTAGGAAATC	TATTGATGC	AGAAGCAGTC	AAGGTAGCC	TaaG.....	
B.b.5		GAAATCACAT	TAcGATTCTC	TTTTTAAGCT	GTTTCCTTAA	ACACTAGTGC	TTAGGAAATC	TATTGAGGC	AGAaCAGTC	AAGGTAGCC	TAGGTTAGT	
B.b.6		GAAATCACAT	TAGGATTCTC	TTTTTAAGCT	GTTTCCTTAA	ACACTAGTGC	TTA.....	...GGAGC	AGAaCAGTC	AAGGTAGCC	Ta.GaTaaG	
B.b.7		GAAATCACAT	TAGGATTCTC	TTTTTAAGCT	GTTTgcTTAA	ACACTAGTGC	TTAGGAAATC	TATTGAGGC	AGAaCAGTC	AAGGTAGCC	TAGGg.TaaG	
B.a.		GAAATCACaA	TgtGATTCTC	TTTTTAAaCt	GTTCCCTTAA	ACACTAGTGC	TTAGGAAATC	TATTGAGC	AGAaCAGTC	AAGGTAGCT	TAGGTTAGG	
B.p.		GAAATgACAT	TAGcATTCTC	TTTTTAAGCT	GTTCCCTTAA	ACACTAGTAc	TTAGGAAATC	TATTGAGGC	AGAaCAGTC	AAGGTAGCC	TAGGTTAGG	
B.b.1	201	GTTAGGCTTA	GGTTAGGCT	TAGGTRAGG	CtTAGGGTAC	GGGTTCCGGG	AGGGGtgcG	GTACGGCCTA	GsaTATGtT	TAGGTTAGG	GTTAGGCTTA	300
B.b.2		GTTAGatTTA	GGTTAGGCT	TAGGTRAGG	gTTeGGGTAC	GGGTTCCGGT	AGGgtTTCG	GTACGgtTGA	GG.....g	TAGGTTAGG	G.TAGtGTTA	
B.b.3		GTTAGGCTTA	GGTTAGGCT	TAGGCTA...	AC	GGGTTCCGGG	.tGGGTTCCG	GTACGGCCTA	GG.....T	atGGsaatGG	GaatGGGTTA	
B.b.4		TTA	GGTTAGGCT	TAGGTRAGG	CtTAGGGTAC	GGGTTCCGGG	AGGGGtTtc	GTACGGCCTA	GGTATGgGt	TAGGTTAGG	GtTAGGCTTA	
B.b.5		GTTAGGCTTA	GGTTAGGag	TAGGTRAGG	GGGTTCCGGG	GGGTTCCGGG	AGGGTTCGG	GTACGGCCTA	GGGTAT...	GG	GTTAGGCTTA	
B.b.6		tTTAGGCTTA	GGGtgaGGCT	TAGGTRAGG	CtTAGGGTAC	GGGTTCCGGG	AGGgtTTCG	GTACGGCCTA	GGGTAT....	..GGGTTAGG	GTTAGGCTTA	
B.b.7		GTTAGGCTTA	GGTTAGGCT	TAGGTRAGG	CtTAGGGTAC	GGGTTCCGGG	AGGgtTTCG	GTACGG...G	GtaTAgGtA	T.GGGTTAGG	GTTAGGCTTA	
B.a.		GTTAGGgTTA	GGGtgaGGt	TgGGGTRAGG	CtTAGGGTAC	GGGTTCCGGG	AGGgtTTCG	GTAGGGCCTA	GGTATGgGt	TAGGTTAGG	tTTAGGCTTA	
B.p.		GTTAGtCTTA	GGGTTAGGCT	TAGGTRAGG	CtTAGGG.cC	cGGTcCGGg	AGGgrTTCG	GTACGrCaTa	GGYAYGgT	TAGGTTAGG	GcTAGGAYYg	
B.b.1	301	GTGTTACGt	TAGGCTCGG	TTTAGGGTAC	GGGTcAGGAT	TAGGTRACGT	GTTAGGTTA	GGGTAGGGT	TAGaCTTAGG	GTACCGCTTA	GGGT.....	400
B.b.2		GTGTTAGGt	TAGGCTCGG	TTTAGGGTAC	GGGTTAGGAT	TAGGTRACGT	GTTAGtGTTA	GgtTAGGGT	TAGGTTAGa	GTACGCaTTA	GGGTAGGGT	
B.b.3		GTGTTAGtT	TAGGCTCGG	TTTAGGGTAC	GGGTTAGGAT	TAGGTRACGT	GTTAGGGTTA	GGGTaGGGT	agGGGTTAGG	GtTaGaGtAc	GcaTTAGGGT	
B.b.4		GgGTTAGtGT	TAGGctTaGg	gTcGGGtTt	aGgTAcGgG	TtaGGattag	GtTAcctTTA	GGGTaGGGT	agGGGTTAGG	GtTaGgGTac	aCtTaAGGtT	
B.b.5		GTGTTAGGt	TAGGCTCGG	TTTAGGGTAC	GGGTTAGGAT	TAGGTRACGT	GTTAGGGTTA	GGGTAGGGT	TAGGTTAGa	GTACCGCaTTA	GGGTAGGGT	
B.b.6		GTGTTAGGt	TAGGCTCGG	TTTAGGGTAC	GGGTTAGGAT	TAGGTRACGT	GTTAGGGTTA	GGGTAGGGT	TAGGTTAGG	GTACCGCTTA	GGGTAGGGT	
B.b.7		GTGTTAGGt	TAGGCTCGG	TTTAGGGTAC	GGGTTAGGAT	TAGGTRACGT	GTTAGGGTTA	GGGTAGGtGT	TAGGTTAGG	G.....		
B.a.		GTGTTAGGt	TAGGCTCGG	TTTAGGGTAC	GGGTTAGGAT	TAGGTRAcTc	GTTAGGTTA	GGGTAGcGGT	TAGGTTAGG	GTACGtGTTA	GGGtgaAGGt	
B.p.		GyGTTAGGgT	TAGGctTCGg	yycAGGGTAC	GGGTcAGGAY	cAGGGTACGc	GtTAcGtTcA	GGGcAGGGT	cAGGccAcGc	GcAcGctTcA	GGGT.....	
B.b.1	401	AGT	TaATAAACCT	ATCCCA..CTC	TAGA	434						
B.b.2		AGT	TaTTAAACCT	ATCCCACTCTC	TAGA							
B.b.3		ta..aggtAGT	TaTTAAACCT	ATCCCACTCTC	TAGA							
B.b.4		AGT	TaTTAA...	TC	TAGA							
B.b.5		AGT	TaTT.....	TC	TAGA							
B.b.6		AtT	TaTTAAACCT	AT.....CTC	TAGA							
B.b.7		AGT	TaTTAAACCT	TC	TAGA							
B.a.		AGT	TaTTAAACCT	ATCACTCTC	TAGA							
B.p.				TC	TAGA							

Fig. 2. Composition of seven sequenced Xbal components of the sei whale (B.b.) and one component of each of the minke (B.a.) and fin (B.p.) whales. Deviations from the consensus sequence are indicated by lower case letters. N:undefined nucleotides; y:c/T; r:A/G.

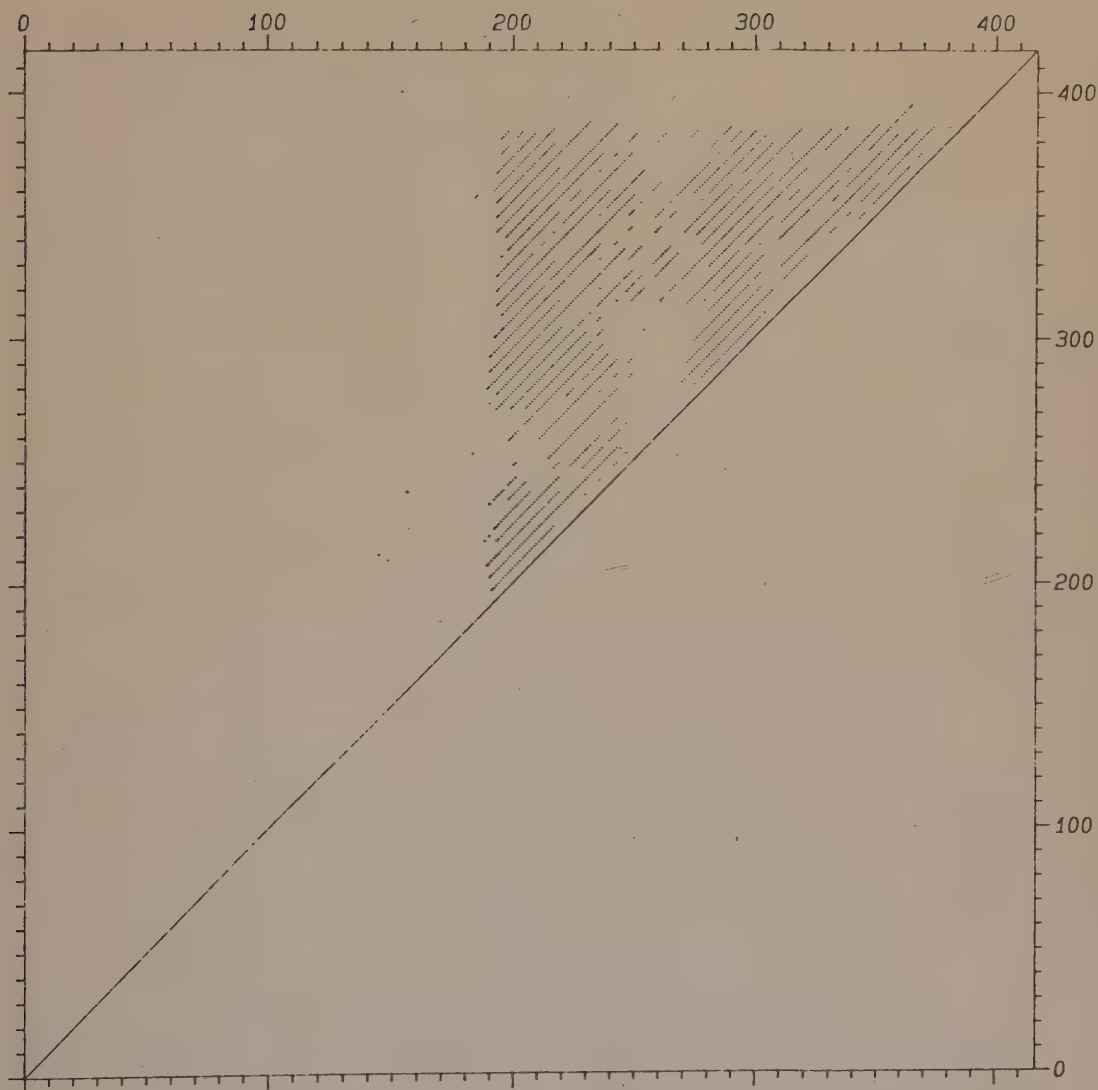


Figure 1. Dotplot of the sei whale Xba I component sequence compared with itself. The plot demonstrates the organization of the sequence in one unique part (positions 1-190 and 390-420) and one repetitive part (positions 190-390) composed of 6 bp repeats.

